Research Report 2016-13

Development of Physical-Empirical models for seasonal prediction of South Korea spring temperature

Vladimir Kryjov

APCC
APEC CLIMATE CENTER

Research Report 2016-13

Development of Physical-Empirical models for seasonal prediction of South Korea spring temperature

Vladimir Kryjov

APCC
APEC CLIMATE CENTER

PREFACE

Physical-empirical models based on physically plausible mechanisms and statistical links complement dynamical model forecasts and enable further improvement of dynamical models by accounting for the empirically established mechanisms. This year, APCC has performed physical-empirical studies in order to increase the reliability of spring seasonal and monthly predictions for East Asia.

This research is focused on the one-month-lead seasonal and monthly prediction of spring temperatures in South Korea. In particular, it investigates the system of forcings and feedbacks affecting the formation of spring East Asia seasonal and monthly air temperature anomalies, as well as the reliable physically plausible and stable in-time lead-lag relationships between the relevant components of the climate system. The obtained knowledge provides a basis for the development of physical-empirical models linking South Korea spring temperature anomalies with the preceding wintertime anomalies of sea surface temperature and circulation using statistical tools.

The main result from the study in 2016 is a set of the physical-empirical models and ready-to-use software tools for the prediction of spring seasonal and monthly temperature anomalies in South Korea, with interpretation in both deterministic and probabilistic terms. These physical-empirical models are mutually complementary with the dynamical models, and possess a prediction skill that exceeds the existing dynamical models. Implementation of the developed methods into the APCC operational practice is expected to provide an essential improvement of our spring seasonal and monthly predictions for South Korea.

I would like to acknowledge the dedicated work of Dr. Vladimir Kryjov in the field of physical-empirical prognostic research at APCC, as well as the APCC staff that supported this work. Also, I would like to recognize that the promising results from the performed research will open the way for further physical-empirical prognostic studies at APCC.

Dr. **Hong-Sang Jung**
President, APEC Climate Center
January 2017

ABSTRACT

Results from the study of the development of physical-empirical models for the prediction of spring South Korean temperatures, i.e., establishing asynchronous physically plausible relationships between spring South Korean temperatures, predictands, preceding climate variables, and predictors, and developing statistical tools for the prediction based on the established relationships, are presented in this report. There are different possible predictors for each spring months, with a transition from extratropical influences dominating the predictions for March and April to tropical influences dominating the predictions for May. The developed models based on multiple regression analysis with a flexible construction of physically plausible predictors have proven their efficiency in tests performed on simulated real-time forecasts for 34 springs (1983–2016). The temporal correlation coefficient (TCC) and mean square skill score (MSSS) of the monthly forecasts vary within 0.47–0.56 and 0.17–0.42, respectively. The ranked probability skill score (RPSS) values of the corresponding probabilistic forecasts are within 0.15–0.30.

The obtained skill implies that the Asia-Pacific Economic Cooperation Climate Center multi-model ensemble (APCC MME) forecasts can be complemented with the forecasts obtained with the developed methods.

CONTENTS

1. INTRODUCTION

The studies of the East Asia climate system are mainly focused on winter and summer conditions. The spring transition from the winter to summer monsoon has not been the subject of careful analyses. The reasons are likely to be the absence of noticeable events (extremes, long-term anomalies, etc.) during the transition season, as well as the complexity of the climate conditions and their causes during the transition season. Winter or summer could be analyzed as a single season, at least in terms of prevailing forcings, while intraseasonal winter or summer variability is associated with the variability of the same forcings within the given season. For example, the winter temperature in South Korea is strongly related to the East Asia Winter Monsoon (EAWM). Enhancement (weakening) of the EAWM is associated with enhancement (weakening) of cold advection to the Korean Peninsula and cooling (warming) over South Korea. This statement is correct at least for the December, January, and February (DJF) season as a whole and for each month separately. We cannot provide a similar generalized cause for spring temperatures. In March, the South Korean temperature and the whole climate system are still under the influence of winter conditions, which is typical for regions with winter ice and snow or those located near seas (e.g., Rigor et al., 2000; Kryjov, 2002; Doblas-Rayes, 2010). Meanwhile, in May, the South Korea temperature is mainly affected by variations in solar irradiance rather than heat advection (Khromov, 1985).

This complexity explains the low skill of dynamical model predictions of Korean spring temperatures. For example, the skill of the Asia-Pacific Economic Cooperation Climate Center (APCC) cross-validated multi-model seasonal predictions in terms of temporal correlation coefficients does not exceed 0.3 (Min et al., 2014). The skill of the spring temperature predictions of World Meteorological Organization Lead Center for Long-Range Forecast Multi-Model Ensemble (WMO LC-LRFMME) is similar (Kim et al., 2016). This explains the interest in statistical methods of forecasting, which often outperform dynamical model predictions in long range forecasting for locations and climatologically (quasi-)homogeneous regions. The term “statistical methods” comprises a variety of approaches. The most common and applicable in long range forecasting are: (1) synoptic-statistical models based on the mostly probable transitions from one synoptic process to another (e.g., Girs, 1960, 1971);

(2) analog methods based on the mostly probable continuation of the observed current process; and (3) physical statistical (also known as physical empirical) methods based on extrapolation into the future of the observed and statistically formalized physically plausible relationships between predictand and predictor(s).

Currently, the most widely used and applied is the third approach, the development of physical statistical (the term accepted in Russia, e.g., Yudin and Mesherskaya, 1977; Ugryumov, 1981; Bagrov et al., 1985) or physical-empirical (the term used, e.g., by Yim et al., 2015; Lee et al., 2013; Wang et al., 2013) models and their application in long-range forecasting. It is worth noting that sometimes physical-empirical models are based on indirect physically plausible relationships or even automatic links (automatic correlation) rather than direct physically plausible relationships; that is, the predictor and predictand could be related to the same physical process while not being physically related to each other and the correlation between them may exceed that between the predictand and the process. It is typical if the process series is noisier than that of the unrelated physical predictor. It should also be noted that higher correlations during the training periods do not provide higher skill in the forecasts.

The main shortcoming of the physical-empirical models is their localization. The predictand is usually a single series representing either a station or a comparatively small area with an area averaging or a principal component (the most popular in the Soviet Union of the 1960s–1980s, e.g., Bagrov, 1959; Aristov et al., 1972), representing a part of the variance of the predictand process. As a rule, the enlargement of the predictand variance requires reduction of the area.

Although there are just a few studies of the spring climate conditions over East Asia that touch on Korea, some information could be extracted from winter and summer studies. First of all, the interannual variability of the EAWM is a regional manifestation of the variability of global processes, particularly, the migration of the polar front jet (Jhun and Lee, 2004; Luo and Zhang, 2015). In addition, polar front jet anomalies in association with spring-early summer precipitation in Eastern China are under discussion (Huang et al., 2014; Li and Zhang, 2014). Ye and Bao (2001) note the association between the low snow cover in the west of Eurasia and positive North Atlantic sea surface temperature (SST) tripole as well as a positive

SST anomaly in the East Asia seas. In support of this linkage, Yim et al. (2015) constructed a set of physical-empirical models for the prediction of Meiyu rainfall in Taiwan based on the Philippine Sea SST and sea level pressure (SLP) anomalies and their (persistent) link to the El Niño Southern Oscillation (ENSO) (Wang et al., 2000), as well as to Atlantic–Pacific influences (Wu et al., 2009; Watanabe, 2004). The Eurasian snow cover influence on the spring-early summer precipitation in East Asia with a focus on China is discussed in a number of studies, e.g., Wu and Kinter (2010); Wu et al. (2014); Wu and Kirtman (2007). The autumn-summer persistent SST, applicable as a predictor in long range forecasting, and the SLP anomaly over the Philippine Sea, governed by the ENSO and Indian Ocean anomalies, are the focus of discussions as well (e.g., Wang et al. 2000; Lau et al., 2005; Yuan et al., 2012).

However, we have not found any study focused on forecasting the Korean spring climate variables or providing a physically plausible basis for such forecasting. This is the main reason for our focus on the development of the physical-empirical models for forecasting Korean spring variables. We focus the study on the development of statistical links based on physically plausible relationships, with the verification tests being performed on real-time retrospective forecasts with no influence from "future" years. The main goal of the study is the improvement of the prediction of spring temperature for East Asia, particularly, South Korea, based on the improvement of our understanding of the East Asia spring climate system. The utilization of this understanding will help in developing prediction schemes and statistical models for the prediction of South Korea spring temperature.

To summarize this discussion, we keep in mind some specific climate features of South Korea, particularly, that the EAWM variability is a dominant feature in the variability of the East Asia climate system in winter (e.g., Jhun and Lee, 2004), that South Korea is surrounded by sea waters, that snow cover is typical in winter (at least in the northern half of South Korea and the mountains) and melts in late winter/early spring, and that an annual cycle of the interannual temperature variance with winter maximum and summer minimum (e.g. Khromov, 1985) is typical for the extratropics. It is logical to suggest that a spring temperature anomaly in Korea will be affected mainly by an anomaly in the preceding winter EAWM intensity. A similar effect was discussed by Kryjov (2002), who has shown that the influence of the anomaly of the winter Arctic oscillation (AO) is traceable in the temperature anomalies along the Barents and Kara sea coast through August due to the memory of the surface properties of the winter conditions when they were formed. A similar effect was also discussed by Kryjov (2004), who has shown that an anomaly in winter zonal circulation is traceable in the northwestern Russia air temperature through June at least due to the persistence of the Baltic Sea water temperature anomaly, which keeps the memory of the winter zonal circulation anomalous impact. Doblas-Reyes (2010) documented the high skill of persistent forecasts of spring temperature over Central and Northern Europe (correlations exceed 0.6), with explanations residing in the "influence of coastal SSTs, the snow albedo, and the latent heat effects of snow melting."

Therefore, based on the general knowledge of the climate of South Korea and surrounding environment, it is logical to suggest that most of the characteristics of a winter anomaly of the EAWM (e.g., EAWM indices, wind characteristics, air temperature, sea heat content, surface temperature, etc.) must be good predictors for South Korean spring temperatures, at least for March-April and, consequently, the seasonal mean, which must be dominated by March-April.

As follows from the author's personal experience of the development of various long-range forecasting methods, predictors with noisy series, even if they provide higher "explained" variance in predictand reconstructions (during the training period), usually fail in forecasts. The least noisy are series of surface and under-surface properties. Keeping in mind the list of potential predictors available, it is reasonable to use as a predictor the SST anomalies in the seas surrounding Korea.

In addition, relationships between a predictand and fixed predictor(s) (point or station variable, index, average over fixed area, etc.) are effective for forecasts during a very limited decadal scale time period. Therefore, it is reasonable to use some constructor of a predictor instead of a fixed prescribed predictor (e.g., Kug et al., 2008; Min et al., 2011; Kryzhov, 2012, Lee et al., 2013). In such an approach, constructor procedures are prescribed rather than their result, a predictor, which provides some flexibility and enhances the effectiveness of the developed methods.

The study is organized as follows. Section 2 briefly describes the data used and the applied methodology of forecasting and verification. Section 3 covers the main research results. Section 3.1 begins with a pre-forecasting analysis of the statistical relationships of spring monthly and seasonal South Korea Temperature (KST). Section 3.2 covers the lead-lag relationships of KST with possible predictors and possible mechanisms for these relationships. Section 3.3 presents the main results, the developed prediction methods for the spring KST anomaly with monthly resolution and verification results. A probabilistic interpretation of the developed deterministic forecasts is given in Section 3.4. Concluding remarks are given in Section 4. A prediction method is detailed in Appendix 1. A special study on the possible influence of the trend in the KST series on the predictions is presented in Appendix 2. Appendix 3 provides a review of the predictors and skill scores.

2. DATA AND METHODOLOGY

2.1 Data

We use temperatures from the University of East Anglia Climatic Research Unit (UEA CRU) dataset of land only temperature and precipitation gridded on a 0.5° × 0.5° mesh (Harris et al., 2014). The KST is represented by the land only temperature time series from the grid points within the trapezoid 38°N-34°N, 125°E-130°E. We use gridded variables from the world recognized climate data center rather than station data because (a) the gridded data are more convenient to use and (b) the time series from the stations and closely located grid points are strongly correlated, with correlation coefficients being well above 0.95.

As predictors, we test variables from various sources. Circulation characteristics are represented by SLP, geopotential heights, and zonal and meridional wind velocity from the NCEP/NCAR Reanalysis-1 (Kalnay et al., 1996). Sea surface properties, SST, and sea ice concentration (SIC) on a 1°×1° grid are from the United Kingdom Met Office (UKMO) Hadley Centre (Rayner et al., 2003). Snow cover (SC) is taken from 20th Century Reanalysis data (Whitaker et al., 2004; Compo et al., 2006; Compo et al., 2011). For illustrations of synoptic phenomena associated with the obtained correlation patterns, we use precipitation and temperature (T2m) from the 20th Century Reanalysis and Reanalysis 1, respectively. The period for the empirical study covers 1958–2016 (59 springs).

2.2 Methodologies

Our forecasting methodology (Appendix 1) is similar to that developed in 2015 and described in APCC Research Report, 2015-09 (Kryjov, 2015) and a paper by Kryjov and Min (2016). It is based on self-tuning physical-empirical models with multiple regression analysis as the statistical tool. It should be emphasized that this method does not imply any prescribed predictor(s) for the entire series of verified forecasts. Instead, a predictor is constructed individually for each forecast based on a unique correlation map estimated individually for the particular forecast without any links to previous forecasts.

The most important condition for verification scores to be reliable is the verification with independent data. Independent data by definition are future observations and their real-time predictions. However, the accumulation of such data in long-range forecasting is very slow. Other data that could be considered independent are historical observations and historical forecasts simulating real-time forecasts. This means that (1) throughout the whole verification period, training years always precede the year of the forecast; and (2) no knowledge of the direct or indirect predictand-predictor(s) relationships (developer's assessments, published results, papers, etc.) from the verification period is used for selection (prescription) of the predictor or construction of the predictor.

Our strategy for verification is simulation of the real-time forecasts, which corresponds to Strategy 2 of Murphy (1988). In this procedure, we operate 26-year samples (series), with 25 years being a training period and the 26th year being the year for the forecast. Throughout the whole verification period, the training years always precede the year of the forecast. Since the whole series is 59 years (1958–2016), a training period size of 25 years allows one to perform 34 predictions for the forecast years from 1983–2016. These results have been verified.

We completely match condition (1) for real-time forecasts for all the forecasts for all months and seasons; that is, training years always precede the year of the forecast. There are no prescribed predictors in use; moreover, even the 25-year length was set quite arbitrarily, with the intent to provide a size for the verification series of not less than 30 years (the 25-year training period is an inheritance from Kryjov (2015) and Kryjov and Min (2016), where a 55-year series was used).

In our case, the basis of a "predictor constructor," that is, the region and predictor variable for which a correlation map is estimated, is prescribed for the whole forecast series. For all predictions from January, the variable (SST) and the area of the predictor constructor (the seas around Korea expanding with time) were selected based only on general considerations (as indicated in the Introduction) without any pre-estimations. However, to argue that our intuitive selection is correct, we have constructed correlation maps (1979–2014, detrended) showing the consistency of the global processes with our choice. Moreover, the correlation maps based on the detrended series are not appropriate for selection of the predictor constructor, since the selection of the original series is needed. Although the January predictor constructors were prescribed regardless of the correlation maps, we do not consider the verification assessments of the forecasts for that period absolutely independent. They are not absolutely dependent as they could be with a prescribed predictor selected for the same period as the forecast verification; however, they are less dependent than those in cross-validated forecasts, since they do not use "future" observations. There are two years, 2015 and 2016, of absolutely independent forecasts. They have not been used in any pre-estimations. Although only two and no reliable numerical assessments are possible, the visual comparison of these forecasts with the observations is quite informative and impressive.

The predictor constructors for the April and May monthly temperature based on February and March are not considered to be independently selected. They were selected based on research results (e.g., Wang et al. 2000; Lau et al., 2005; Yuan et al., 2012; Huang et al., 2014; Li and Zhang, 2014) supported by the correlation maps. However, the verification assessments based on these forecasts are more independent than those based on cross-validated forecasts. For the 2015 and 2016 forecasts, they are absolutely independent.

The performance of the forecasts is assessed by a mean square skill score (MSSS) recommended by WMO (2012), with reference forecasts being climatological (Murphy, 1988), and a temporal unadjusted correlation (congruence) coefficient (TCC), also known as an anomaly correlation coefficient. Detailed explanations of the forecasting and verification procedures are explained in Appendix 1 and Appendix 2 of APCC Research Report, 2015-09 (Kryjov, 2015). The probabilistic forecasts are assessed with the rank probability skill score (RPSS) and relative operating characteristic score (ROCS) equal to the area under the ROC curve (WMO, 2002).

The MSSS and RPSS vary from $-\infty$ to +1 and its probability distribution function is not tabulated, therefore the statistical significance of MSSS, RPSS, TCC, and ROCS was assessed using a bootstrap method with 1000 Monte Carlo trials accounting for serial correlation (Wilks, 1995; 1997).

Confidence intervals of the correlation coefficients have been estimated based on a Fisher z-transformation (Wilks, 1995), accounting for serial correlation

(Bretherton et al., 1999). The average values of the correlation coefficients have been estimated using the Fisher z-transformed coefficients with further inverse transformation.

A non-related to forecast supporting correlation analysis (correlation maps, correlations between fixed series, etc.) was performed on the series from 1979–2014 (36 years). The effective sample (series) size accounting for serial correlation (Bagrov, 1969; Bretherton et al., 1999) for most of the grid points is above 33 years, the corresponding number of degrees of freedom is above 31, and the corresponding correlation threshold value for the 5% significance level is approximately 0.35.

Additional composite analysis was performed on the links between the January, February, and March fields of SST, East Asia air temperature, and surface wind. This analysis completely supports the results from the correlation analysis, so it is not shown in the report as the presentation of these results requires dozens of pages with pictures.

As a comparison, we performed a verification based on cross-validated forecasts. Consistent with the real-time forecast simulations, we applied a nine-year-leave-out cross-validation on the series from 1983–2016; that is, for 34 forecasts. Quite expectedly, the verification scores obtained in the cross-validation are much higher than those obtained in the simulations of real-time forecasts. Since verification based on cross-validated forecasts is inappropriate for any method of forecasting that implies that tuning is performed on the series with non-zero serial correlations, we do not discuss and do not show a skill for the cross-validated forecasts in the report so as not to confuse the readers.

3. RESEARCH RESULTS

3.1 South Korea Spring Temperature. Peculuarities of the Time Series

Spring is a transition season from winter, when South Korean temperatures are mainly affected by heat advection anomalies associated with the EAWM (Jhun and Lee, 2004), to summer, when temperature anomalies depend mainly upon solar irradiance anomalies rather than direct heat advection (Khromov, 1985).

The time series of the South Korean spring seasonal mean and monthly temperatures are shown in Figure 3.1.1. The most notable feature is a decrease in monthly variance from March to May. Indeed, the March KST standard deviation is 1.2 K, while April and May standard deviations are 1.0 K and 0.6 K, respectively. This ratio is typical, within a 95% confidence interval, for the whole period of analysis as well as for its halves. Taking into account that the correlation between March and April KST is 0.38 and the correlation between April and May KST is 0.44 (Table 3.1.1), it becomes clear that the variability of the seasonal mean spring KST is dominated by the March temperature. Therefore, the success of the prediction of the seasonal mean March, April, and May (MAM) KST depends, to a large degree, upon the success of the March KST prediction. Predictors selected for the March KST are likely to be appropriate for the MAM KST. It should be noted that the above consideration for the predictors of MAM KST is quite formal and touches on only the numerical scores of prediction skill. It does not refer to the physical plausibility of the relationships lying beyond the predictions. Particularly, a low correlation between the March and May KST series (0.03) implies a difference in mechanisms for the formation of monthly anomalies between the early and late spring KST anomalies as well as a difference in predictors.

The generalized lag correlations between the South Korean monthly mean temperature series are shown in Table 3.1.1. The one-month lead correlation between the March and January temperature series is 0.533. It is a reasonable value for testing the January temperature as a predictor for the March temperature. An increase in the correlation from 0.418 (January-February) to 0.533 (January-March) may be typical (the difference is within the 95% confidence intervals for both correlations), or it may reflect amplification of the winter conditions' delayed impact during the period of snow/ice melt, similar to the amplification for the Barents and Kara sea coastal stations described by Kryjov (2002).

Table 3.1.1 Correlation coefficients between the series of monthly KST, 1979–2013, detrended. Coefficients that are significant at the 95% and higher confidence level are typed in bold.

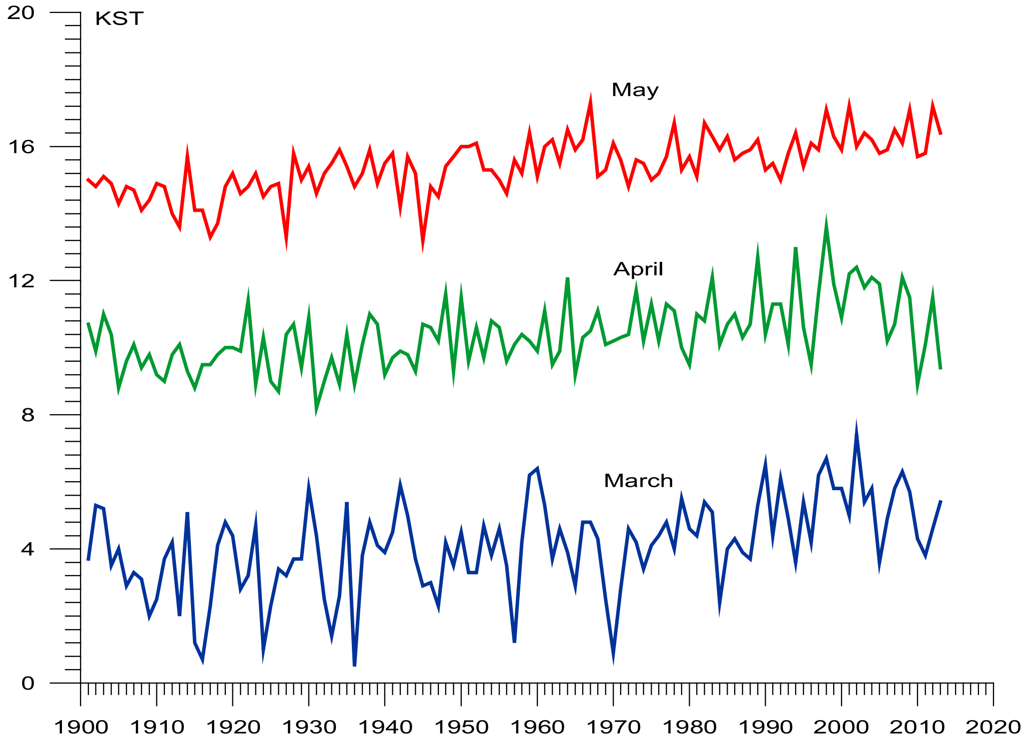
KST/KST	Jan	Feb	Mar	Apr	May
Jan	**1.00**	**0.42**	**0.53**	0.31	0.03
Feb		**1.00**	**0.58**	0.23	0.08
Mar			**1.00**	**0.36**	0.03
Apr				**1.00**	**0.44**

Another peculiarity of the series shown in Figure 3.1.1 is the distinct positive trend. Indeed, the trend assessments, presented in Table 3.1.2 and Figure 3.1.2, support this, with the seasonal and all monthly trends being significant at the 5% level for the 55-year series (1958–2012) and on some series within the 25-year windows. These trends essentially complicate the development of physical empirical models, since relationships between both original and detrended predictands and predictors are to be analyzed. In addition, the trends complicate the verification assessments of the predictions because they require the consideration of the trend forecast rather than a simple climatological one. The trend related issues are detailed in Appendix 1.

Table 3.1.2 Linear trend characteristics of South Korea temperature series 1958–2012, trend value and Student's *t* statistic.

	Trend K/55 years	Student's *t* statistic
March	1.34	2.56
April	1.22	2.88
May	0.62	2.35
MAM	1.06	3.69

(a)



(b)

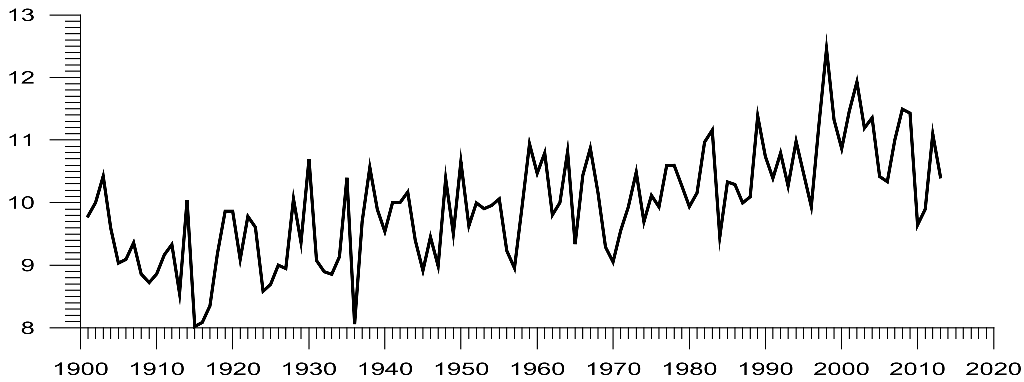
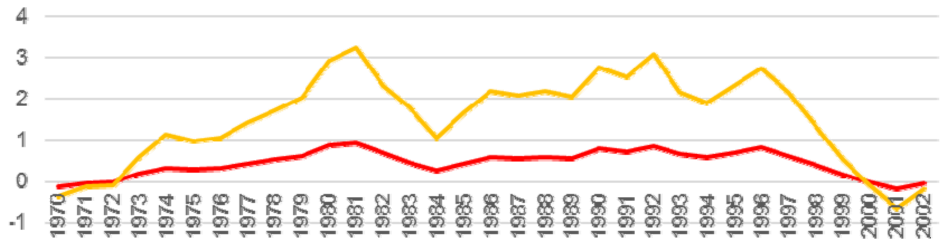
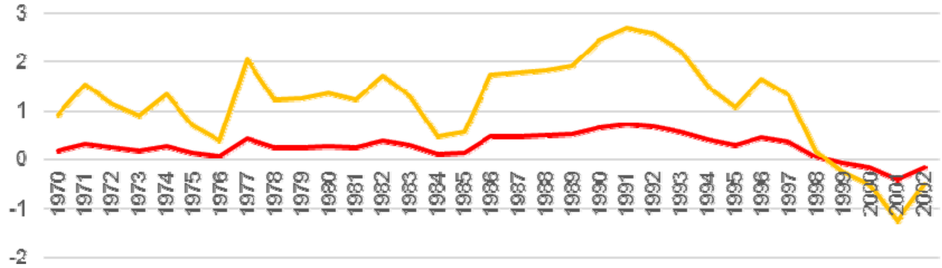


Figure 3.1.1 Time series of MAM (a) monthly and (b) seasonal mean air temperature averaged over South Korea (land only temperature from the trapezoid 38° N-34° N, 125° E-130° E). Years as of January.

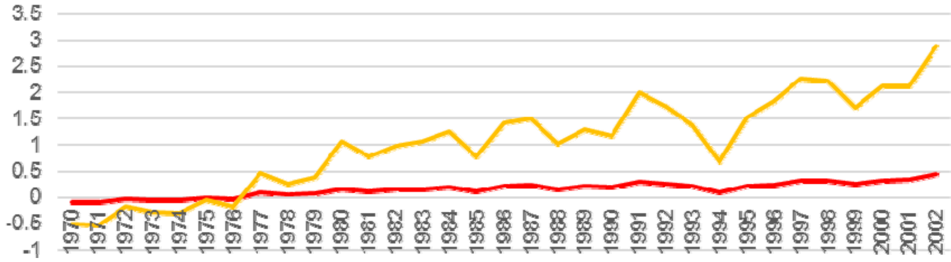
(a)



(b)



(c)



(d)

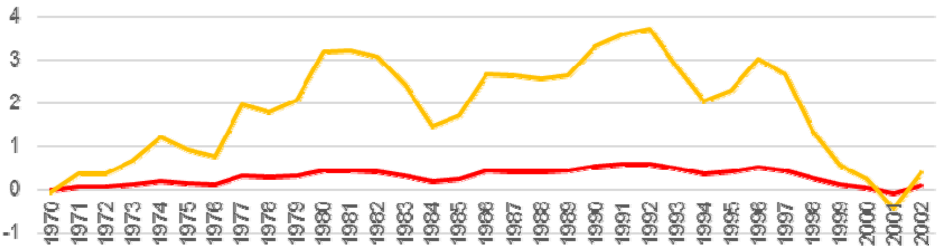


Figure 3.1.2 Trend assessments in the 25-year sliding windows. Red line-trend as a linear regression coefficient, K/decade; yellow line-Student's statistic *t*. Trends in the 25-year windows are significant at the 5% level when *t* exceeds approximately 2. (a) March, (b) April, (c) May, and (d) MAM.

3.2 Pre-Prediction Analysis

3.2.1 Lag Correlations in South Korean Winter-Spring Temperature

The following analysis is performed on the series of temperature fields from 1979–2014 (36 years). The effective sample (series) size accounting for serial correlation (Bagrov, 1969; Bretherton et al., 1999) for most of the grid points is above 33 years, the corresponding number of degrees of freedom is above 31, and the corresponding correlation threshold value for the 5% significance level is approximately 0.35.

A set of correlation maps (Figure 3.2.1) shows the relationships between the grid point temperature series in a set of successive months, i.e., January and February, February and March, etc. These correlations could also be determined as one-month lag/lead correlations and be used for a zero lead forecast of monthly temperature. The level of correlations from January to February and from February to March is mainly above 0.4, i.e., significant at least at the 5% level. However, from March to April and from April to May, the correlations steadily decrease. Correlations exceeding 0.5 over South Korea appear only between the March and February series (Fig. 3.2.1b), with several coastal stations featuring correlations exceeding 0.6. In general, these correlation maps indicate the location of the highest correlations at the western and eastern coastal strips in late winter/early spring and by the southern coastal region in late spring. This is easily understandable, accounting for the prevalence of northwesterly wind in winter/early spring and southeasterly winds in late spring. In combination with the background advection, there exists a night-breeze-like circulation, which adds warming over the coasts in winter/early spring due to vertical air movements, that ascend above the sea associated with moist adiabatic lapse rate cooling and descend over the land associated with dry adiabatic warming.

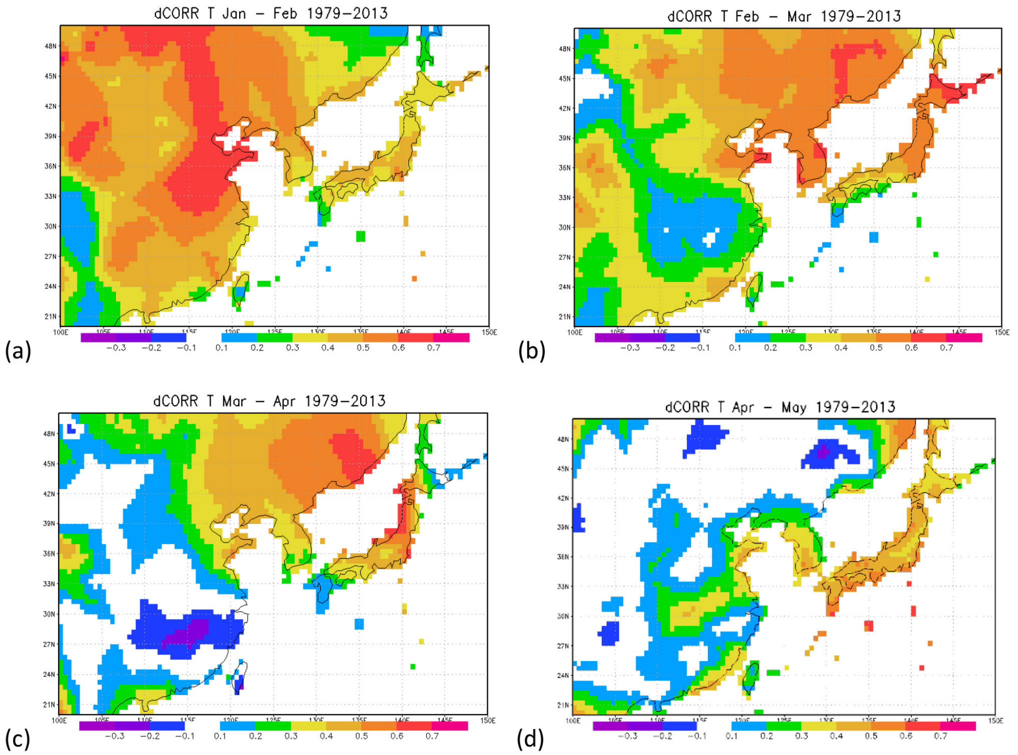


Figure 3.2.1 Successive maps of grid point temperature correlations (a) Jan–Feb, (b) Feb–Mar, (c) Mar–Apr, and (d) Apr–May.

The two-month lag correlation maps for each grid point correlation coefficient between the 36-year series of temperature observed in, e.g., January and March, are shown in Fig. 3.2.2. These two-month lag correlations correspond to one-month lead forecasts, e.g., the forecast of March temperature based on January temperature. For the January-March map, most of the grid points have correlations exceeding 0.4, with correlations exceeding 0.5 for the South Korean west coast and exceeding 0.6 for the northern South Korean grid points (Fig. 3.2.2a). This map shows the sources of the correlation of 0.53 between January and March KST (Table 3.1.1). The correlation pattern between January and the seasonal mean MAM temperature series (Fig. 3.2.2d) closely resembles that of January-March, with the general level of correlations being some 0.1 lower than that of January-March. This similarity in the patterns confirms the thesis regarding domination of the March temperature variability in the seasonal mean. The two-month lag correlations decay with seasonal warming, and correlations between the March and May series are significant at the 5% level only along the east coast of South Korea.

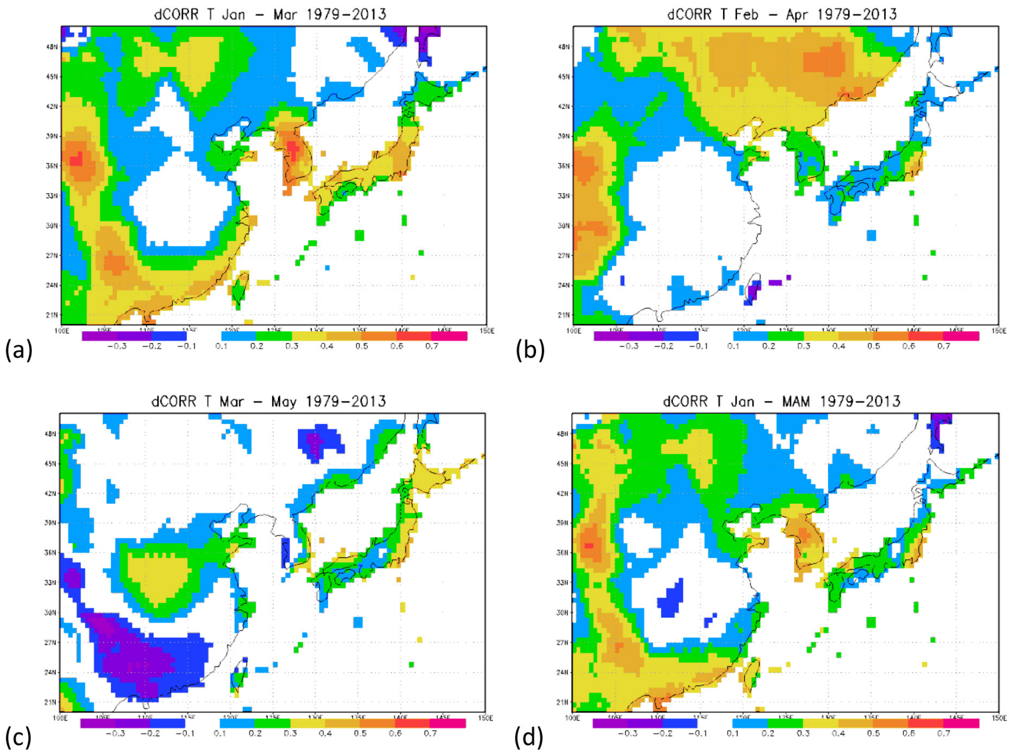


Figure 3.2.2 Successive maps of grid point temperature correlations with one-month lead (a) Jan-Mar, (b) Feb-Apr, and (c) Mar-May; (d) grid point temperature correlations with one-month lead Jan-MAM.

Analysis of the correlation maps in Figs. 3.2.1 and 3.2.2 shows that the influence of the local winter conditions is easily traceable only in early spring, particularly March, leaving consideration of the MAM anomalies to be dominated by the March anomalies. These maps indicate that the most probable agents providing this influence are the anomalous heat content of the waters surrounding Korea and anomalous snow amount on the land.

3.2.2 Mechanisms Providing Lag Correlations

Late winter to early spring, the lead time for prediction for March based on January conditions, is the season for climatological air warming from -2.0 °C to 5.0 °C in Korea. This climatological process governed by the seasonal increase of solar irradiance is climatologically slowed by the water phase transition from ice through water to vapor and is climatologically accelerated by the surrounding sea waters that heat the air coming to Korea from the Asian continent over these seas. The SST decreases from 12.1 °C in January to 11 °C in February-March.

Therefore, the March temperature anomaly is controlled by the initial conditions, the January KST anomaly, and anomalies in the conditions for seasonal warming deceleration and acceleration, which can be considered as boundary conditions. This scheme is realizable in dynamic model simulations. However, it is hardly realizable in a statistical prediction model, mainly because of the series noise and the non-deterministic relationships between the variables in the statistical model suggesting some uncertainty. It is more reasonable to construct a physically plausible statistical model based on co-variability of the relevant variables on an interannual time scale.

As mentioned previously, persistence of the anomalous heat content of the waters surrounding South Korea provides the impact of the winter conditions on the early spring conditions. The persistence of anomalous heat content to a large degree is reflected in the persistence of surface temperature anomalies, which is well illustrated by the values in Table 3.2.1. For KST, even with the KSST (and other) support for winter, the lag for a reasonable minimum forecast correlation does not exceed one to two months (Table 3.1.1). Meanwhile, due to the high persistence of the waters surrounding Korea, the trail of the winter KSST anomalies is well recognizable throughout the whole spring at least, with the January-to-March correlation exceeding 0.7 (Table 3.2.1).

Table 3.2.1 Correlation coefficients between the series of monthly sea surface temperatures averaged over the seas surrounding South Korea (30–45°N; 120–130°E), 1979–2013, detrended. Coefficients significant at the 95% and higher confidence levels are typed in bold.

KSST	Jan	Feb	Mar	Apr	May
Jan	**1.00**	**0.88**	**0.73**	**0.70**	**0.50**
Feb		**1.00**	**0.87**	**0.79**	**0.57**
Mar			**1.00**	**0.85**	**0.65**
Apr				**1.00**	**0.85**

Anomalies in both winter KST and KSST are representative of the intensity (anomaly) of the EAWM. The intensity of the EAWM is appropriately represented by the index of EAWM (EAWMI) suggested by Jhun and Lee (2004), characterizing the shift in the position of the polar front jet over East Asia. Correlation of the December-January EAWMI with the January KST is -0.63 and -0.55 with the January KSST (Table 3.2.2).

Table 3.2.2 Correlation coefficients between the series of DJ mean EAWM index of Jhun and Lee (2004) and monthly mean KSST and KST, 1979–2013, detrended. Coefficients significant at the 95% and higher confidence level are typed in bold.

DJ EAWMI	Jan	Feb	Mar	Apr	May
KSST	**−0.55**	**−0.62**	**−0.69**	**−0.54**	−0.29
KST	**−0.63**	**−0.56**	**−0.50**	−0.14	0.04

Correlations of the DJ EAWMI with KSST increase from January to March, reflecting stronger cooling of the seas around Korea during the winters with a strong EAWM and when an early spring follows. It should also be noted that there is a strong correlation between the January and February monsoon indices (0.67). Meanwhile, the DJ EAWMI correlations with KST slightly weaken with time, remaining significant (5%) through March. These lag correlations reflect a lag effect from the anomalous (due to the EAWM anomaly) formation of the winter SST field and associated KSST anomaly with the following concurrent influence on the KST anomaly. It is supported by the high concurrent correlations between KST and KSST (Table 3.2.3).

Table 3.2.3 Correlation coefficients between the series of monthly KST (columns) and KSST (rows), 1979–2013, detrended. Coefficients significant at the 95% and higher confidence level are typed in bold.

KSST/KST	Jan	Feb	Mar	Apr	May
Jan	**0.67**	0.26	**0.39**	0.27	0.09
Feb	**0.72**	**0.58**	**0.59**	0.32	0.05
Mar	**0.70**	**0.64**	**0.76**	0.29	0.09
Apr	**0.56**	**0.52**	**0.69**	**0.65**	0.23
May	**0.43**	**0.43**	**0.64**	**0.77**	**0.46**

The DJ EAWMI strongly correlates with the March KSST (–0.69) and March KST (–0.50), and these two correlate with each other with a correlation coefficient of 0.76. There are several mechanisms supporting these relationships implicitly, and they can be analyzed on the basis of the anomalously strong winter EAWM. The first mechanism is the direct impact of the enhanced northwesterly winds on the sea surface, upwelling along the Asian coast, with enhanced storminess and an enlarged mixed layer, which results in cooling of the sea surface and a lower SST. At the same time, a stronger winter EAWM results in a lower temperature (–0.35), enlarged snow accumulations over Eastern China (0.35), and a lower air temperature over Eastern China in February and March during the snow melt (0.41). Even although the sensible and latent heat fluxes are above normal due to increased wind velocity and increased difference between water and air temperature and humidity values, both the sea surface and the air above it remain colder than normal. Therefore, this process also results in a lower SST. As a result, the colder than normal air comes to Korea, providing a negative temperature anomaly. An anomalously weak EAWM leads to a positive temperature anomaly over Korea in early spring.

Another mechanism, which leads to the same results, underlies the temperature lag correlations in northern South Korea (Fig. 3.2.2a). The strong winter EAWM leads to a decrease of temperature (–0.63) and enhanced January snow accumulation in South Korea (corr = 0.44), with the correlation between January temperature and snow cover being –0.52, consistent with the results of Clark et al. (1999) that document that snow accumulation under temperatures close to zero depends more on temperature than precipitation. The snow accumulated by March strongly affects the March temperature (corr = -0.56), resulting in negative temperature anomalies during snow melt. Later, tracking soil moisture (Wu et al., 2014) with the heat of absorption from two water phase transitions provides the effective deceleration of the spring warming (Groisman et al., 1994; Rigor et al., 2002; Kryjov, 2002). The weak EAWM leads to a weak deceleration and positive temperature anomaly.

3.3 Prediction of the Spring Temperature Anomalies

As noted previously, the seasonal mean KST anomaly is, to a large degree, defined by the March KST anomaly. Therefore, predictors selected for the March KST could be used for the prediction of the seasonal mean KST without an essential degradation in the prediction skill. However, spring is a transition season from winter to summer, which implies transition in the factors mainly influencing KST, from circulation to insolation. Therefore, a prediction method has been developed for each spring month separately.

3.3.1 March

3.3.1.1 Concurrent Relationships of March Temperature with Circulation and Surface Properties

Maps of correlation coefficients (detrended, 1979-2014) between the March KST and climate variables representing the concurrent March climate system are shown in Fig. 3.3.1.1. A positive temperature anomaly over South Korea is associated (we use this term because of the analysis of concurrent relationships) with a northward shift of the polar frontal zone to Northern Eurasia (Scandinavia-Kamchatka belt) from the Mediterranean Sea-mid-China belt, which is reflected in the SLP plus a tendency to a positive North Atlantic Oscillation (NAO) phase, precipitation, and temperature, with reduced snow cover in Europe, reduced snow extent throughout Eurasia, and a positive SST anomaly along the European coast (result of enhanced westerly winds) and in the seas surrounding Korea (result of the weakened EAWM).

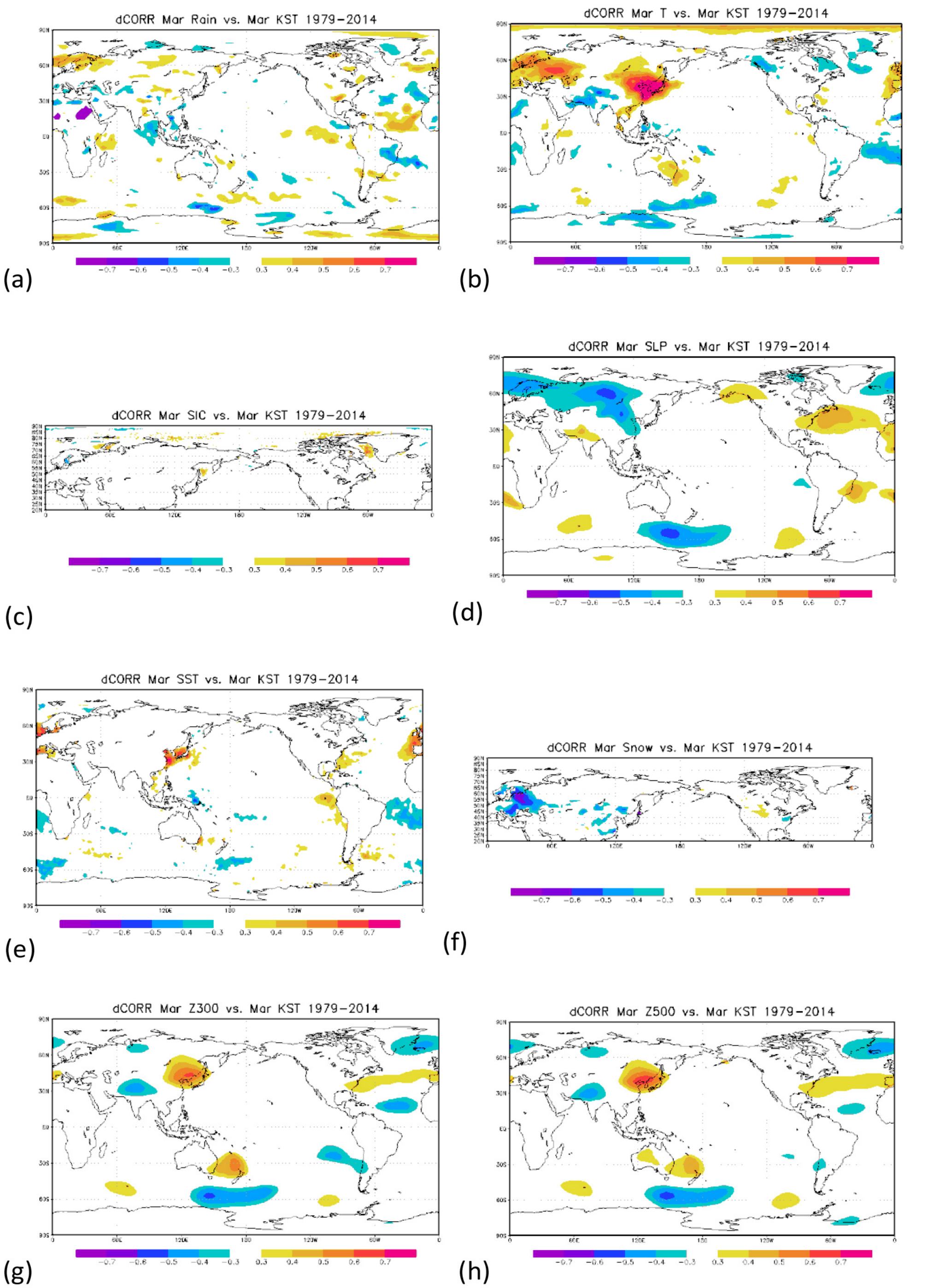


Figure 3.3.1.1 Concurrent correlations between March KST and (a) March precipitation, (b) temperature, (c) sea ice concentration, (d) sea level pressure, (e) sea surface temperature, (f) snow cover, (g) Z300, and (h) Z500, 1979-2014, detrended.

All these phenomena indicate the enhancement of zonal circulation in the middle to high latitudes with a northward shift of the main climatological fronts. The North Atlantic patterns of Z500 and Z300 indicate a positive phase of the NAO. A strongly pronounced pattern of positive correlations between KST and Z500 and Z300 over Korea indicates a weakening of the East Asia trough (northward shift of the jet similar to winter) combined with a positive temperature anomaly in the layer of troposphere below Z500 and Z300.

3.3.1.2 Lag Relationships of March Temperature with Preceding January Circulation and Surface Properties

Two-month lag correlation maps (corresponding to a one-month lead forecasts) between March KST and January variables considered potential predictors of the March KST are shown in Fig. 3.3.1.2.

The two-month lag correlation maps are consistent with each other and qualitatively resemble those showing concurrent relationships. A northward shift of the polar frontal zone, weakening of the Ural high and the northern Siberian high, and a negative polarity of the EAWM upper troposphere high over Korea are apparent from Figs. 3.3.1.2.g and 3.3.1.2.h and imply an anticyclonic shift in zonal wind, matching the inverted cyclonic shift of EAWM of Jhun and Lee, 2004 (see their Fig. 2). The relationships presented in Figure 3.3.1.2 suggest that we may consider any January (and earlier) variable or derivative representing the polarity and intensity of the January EAWM as a potential one-month lead predictor for the March KST. Although correlations of the predictand with atmospheric potential predictors are usually higher on the dependent series (training period) than those with surface potential predictors, experience shows that atmospheric series are noisier than surface series; and, therefore, for independent forecasts, it is more reasonable to use surface series (e.g., Ugryumov, 1981). In addition, a physical empirical long range forecasting is based on physically plausible lead/lag relationships, which implies that a signal is transferred from a predictor (or related variable) to a predictand (or related variable) via some persisting (slowly changing) environmental process and/or via an obvious succession of the processes.

From general considerations, it is reasonable to use January KSST (sea surface temperature in the seas surrounding Korea) as a predictor because it is concurrently and asynchronously related to the winter EAWM. The sea mixed layer, the temperature of which is represented by SST, is able to keep the memory of the January formation of its anomalies through mid-spring (Table 3.2.2); and through the sensible and latent heat fluxes, it is able to heat the air that comes to Korea from the Asian Continent with the westerly and northwesterly winds. It should be noted that the heating of the air is associated with the cooling of the sea, with a stronger EAWM causing a stronger cooling (correlation above the 90% (95%) significance level for January (February)) EAWMI.

It should be emphasized that the January KSST represents the early winter EAWM. It is strongly correlated with other EAWM characteristics, particularly, the EAWM indices, EA temperature, KST, Siberian high, EA snow cover, etc. This means that neither of these characteristics is reasonable to use as an additional predictor in a multiple correlation prediction scheme. Therefore, we use a one predictor regression.

The prediction scheme for the March KST anomaly using the January KSST anomaly as a predictor is based on physically plausible mechanisms. The January negative (positive) KSST anomaly is a result of a strong (weak) early to mid-winter EAWM. This water temperature anomaly persists into March. In late winter, it is enhanced (weakened) by the February EAWM, which strongly succeeds the January EAWM (corr = 0.67). In March, the KSST anomaly is negative (positive) after a strong (weak) EAWM winter. In March, following a winter with a strong (weak) EAWM, wind from any sea direction brings (in terms of anomalies) cooling (warming) to Korea depending upon the sign of the SST anomaly that persists from January.

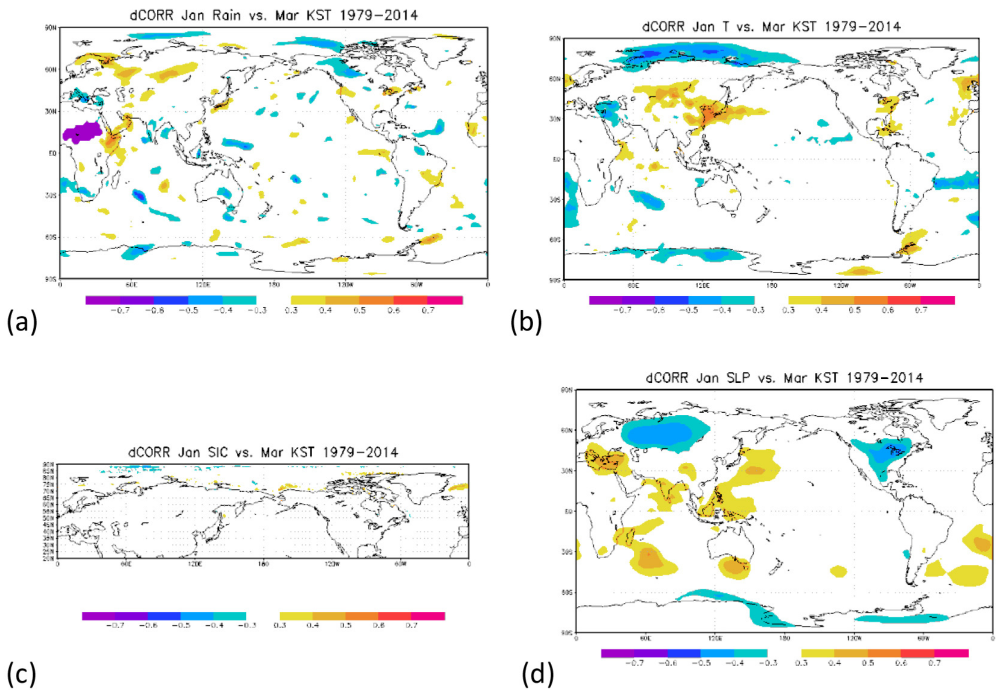
Generally, we appoint SST as a predictor variable, and for the domain for predictor construction, we enclose the seas surrounding Korea at 30–45°N, 110–140°E. A map averaging over 34 training period correlation maps between the January wind components and constructed predictors for March KST is shown in Fig. 3.3.1.3. The main feature demonstrated by this map is an anomalous southeast wind over Korea and the surrounding seas. It indicates some weakening of the northwest wind associated with the EAWM and weakening of the EAWM as a whole, as shown in Fig. 3.3.1.2.

3.3.1.3 Prediction of March Temperature

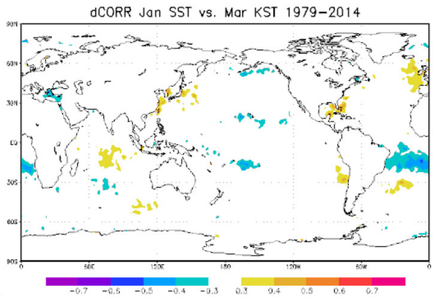
We performed 34 forecasts of the March temperature anomalies for the years 1983–2016. The statistical prediction model for these forecasts is a single predictor regression equation of the form:

$$KST = k\ KSP \quad \text{(Eq. 3.3.1.3.1)},$$

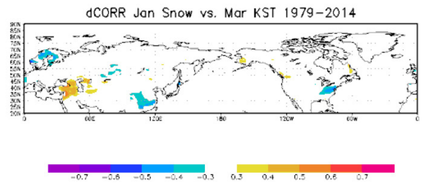
where KST is the March KST anomaly, KSP (Korea Seas Predictor) is a predictor constructed from the January SST anomalies (Fig. 3.3.1.4a) from the trapezoid 45–30°N, 110–140°E, and k is the regression coefficient.



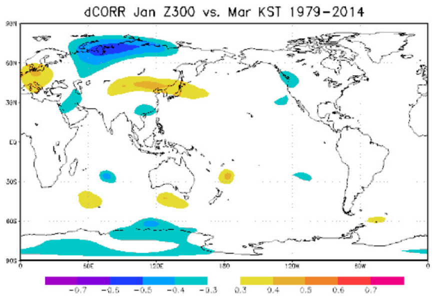
(continued on the next page)



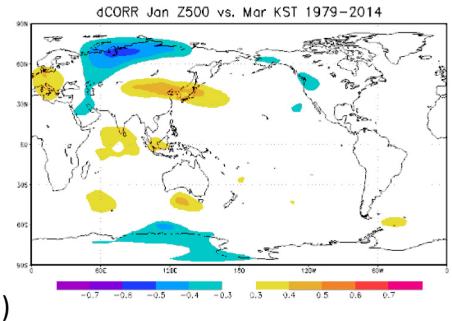
(e)



(f)



(g)



(h)

Figure 3.3.1.2 Two-month lag correlations between the March KST and (a) January precipitation, (b) temperature, (c) sea ice concentration, (d) sea level pressure, (e) sea surface temperature, (f) snow cover, (g) Z300, and (h) Z500, 1979-2014, detrended.

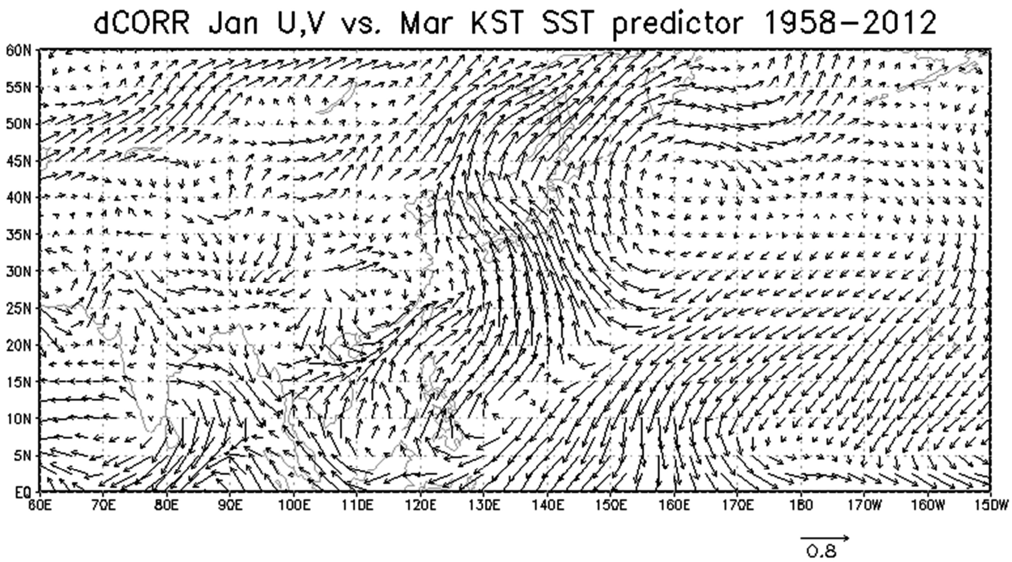
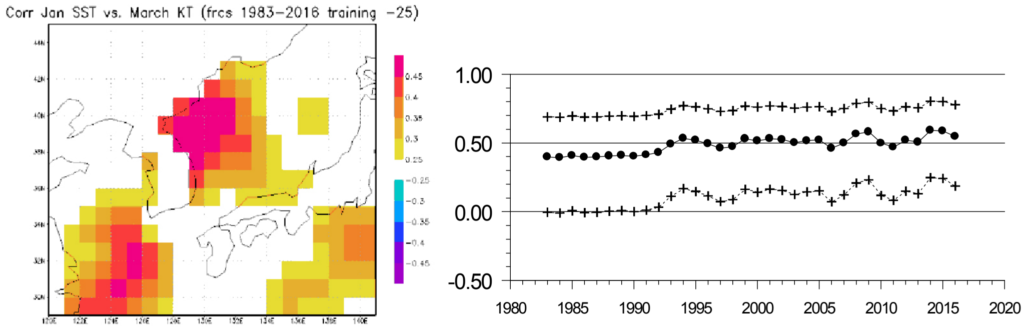


Figure 3.3.1.3 Correlation map averaged over 34 training period correlation maps between the January wind components and constructed predictors for the March KST.

The normalized regression coefficient for each forecast is shown in Fig. 3.3.1.4b. For a single predictor regression, it is equal to the corresponding correlation coefficient. Each normalized regression (correlation) coefficient value is plotted in the forecast year, with correlated series covering the preceding 25-year training period. The relationships between KST and the constructed predictors is quite stable over time, with correlations for the training periods starting with 1968-1992 mainly varying between 0.5-0.6. Correlations for previous training periods are about 0.4. The cause of this 0.1-0.2 step is not obvious. It could be a change in data quality, or it could be due to some variation in the relationships. A possible climate shift in the 1990s is discussed in a number of studies (e.g., Yim et al., 2013). Certainly, this period of forecasts (1983-1992) deteriorates the final verification scores assessed over the whole period of forecasts (1983-2016). Nevertheless, due to the independence of the forecasts and, particularly, the independence of the predictors, constructed individually for each forecast, the predictors and forecasts from this earlier period do not affect the predictors and forecasts of the later period.



(a) (b)

Figure 3.3.1.4 (a) The average of 34 correlation maps used for construction of the predictors. (b) Correlation coefficients (circles) between the predictand and constructed predictors. Correlation coefficient values are plotted in the forecast years, with correlated series covering the preceding 25-year training periods. Crosses denote the 95% confidence intervals. Real-time forecast scheme.

The series of observed and predicted anomalies of the March KST are shown in Fig. 3.3.1.5. The verification scores of the predictions are shown in Table 3.3.1.1. The forecast verification scheme simulating real-time forecasts results in PCC = 0.45, TCC = 0.56, and MSSS = 0.25. This definitely outperforms the model forecasts.

A comparison with other statistical forecasts is impossible because no such verification results have been published in the available journals so far.

It is worth a reminder that although the forecast scheme arose from general considerations rather than any pre-assessments and the verification is based on the simulation of real-time forecasts rather than cross-validation, there may be doubts about the independence of the forecasts up to 2014. Nevertheless, the forecasts for 2015 and 2016 (see Fig. 3.3.1.5) are absolutely independent, and they can provide some basis for the independent assessment of the developed forecast method. Therefore, the developed physical empirical model provides a quite successful prediction of the March temperature averaged over South Korea.

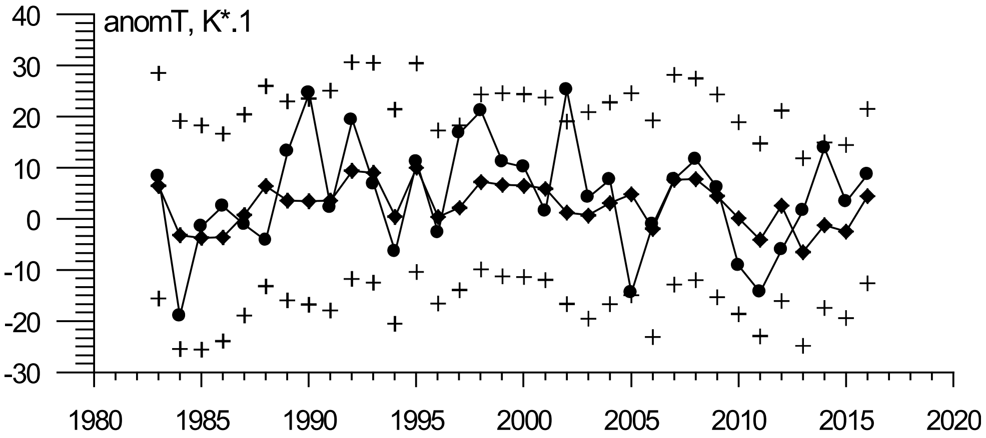


Figure 3.3.1.5 Series of observed (circles) and predicted (rhombs) anomalies of the March KST. Crosses denote the 95% confidence intervals of the forecasts. Real-time forecast verification scheme.

Table 3.3.1.1 Skill scores of March forecasts.

PCC	TCC	MSSS
0.45	0.56	0.25

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

3.3.2. April

3.3.2.1 Concurrent Relationships of April Temperature with Circulation and Surface Properties

In April, a positive (negative) anomaly of temperature in South Korea is associated with an enhancement (weakening) of the frontal zone oriented from Central Europe to the Chukotka Peninsula, which is clearly pronounced in the consistent anomalies of circulation at Z500 and Z300, SLP, anomalies of precipitation, snow cover in Manchuria, and temperature in East Asia. The SST anomalies in the North Atlantic match the positive phase of the NAO. Local East Asia peculiarities include an SLP trough oriented from the Baikal Lake to southeast China and an anomaly of SST in the seas surrounding Korea, consistent with both the SLP anomaly and the high persistence of the sea temperature anomaly from March.

In terms of macro-circulation, the positive temperature anomaly in Korea in April is associated with the intensification and northward shift of the polar jet stream, also called the polar front, storm-tracks, etc. (the positive phase of the AO, NAO (Hall et al., 2015) etc.), with a corresponding weakening of the East Asia trough. In addition, it is associated with a positive KSST anomaly. However, the persistence of the winter sea water temperature anomaly decays with time (Table 3.2.3), and the correlation between the DJ EAWMI and April KSST is only 0.54 versus 0.69 for the March KSST. In April, a local spring wind anomaly becomes important and is reflected in correlations between SLP and KST (Fig. 3.3.2.1d), while a similar correlation map for March (Fig. 3.3.1.1d) reflects a general EAWM related pattern rather than a local one.

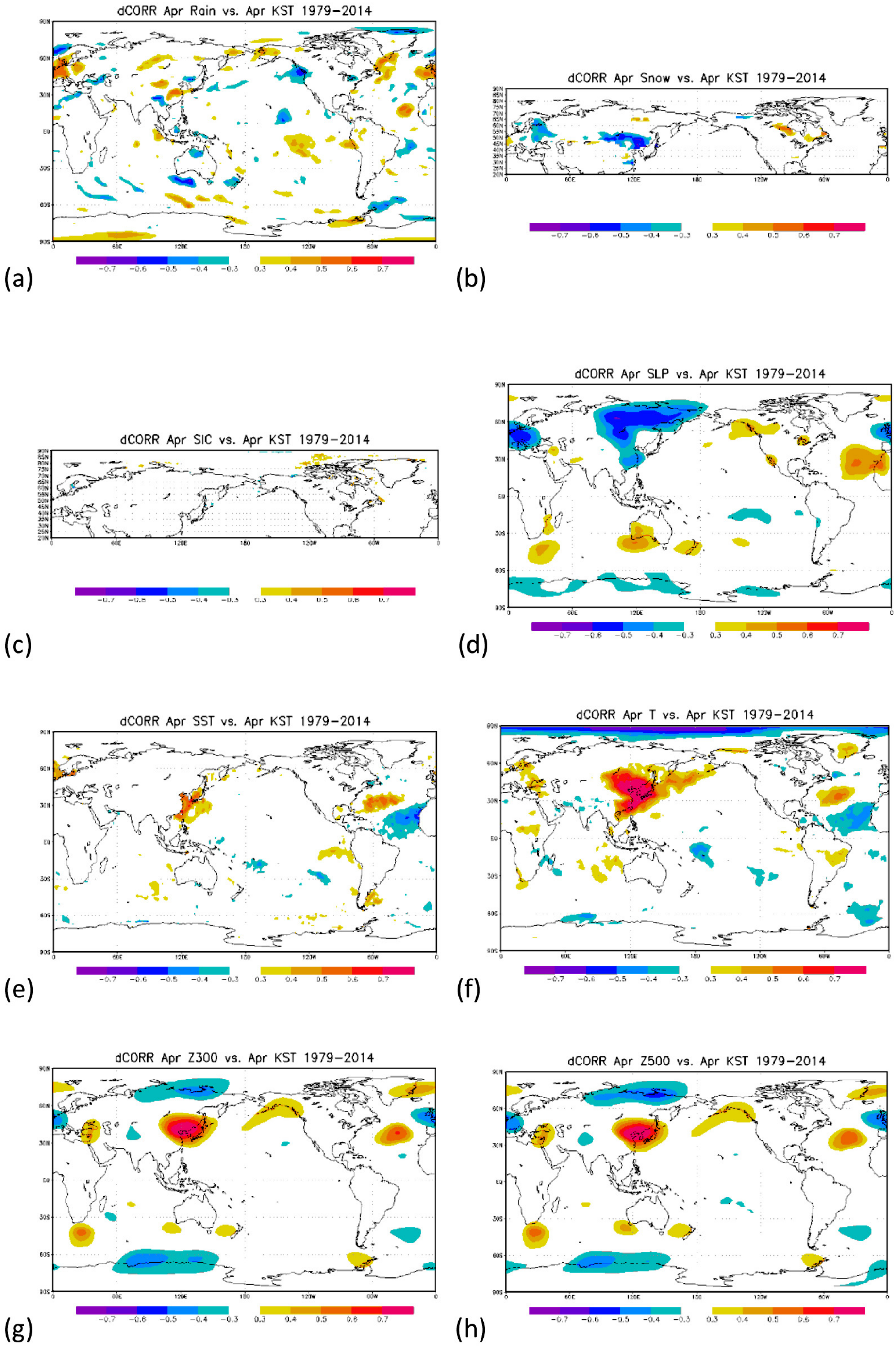
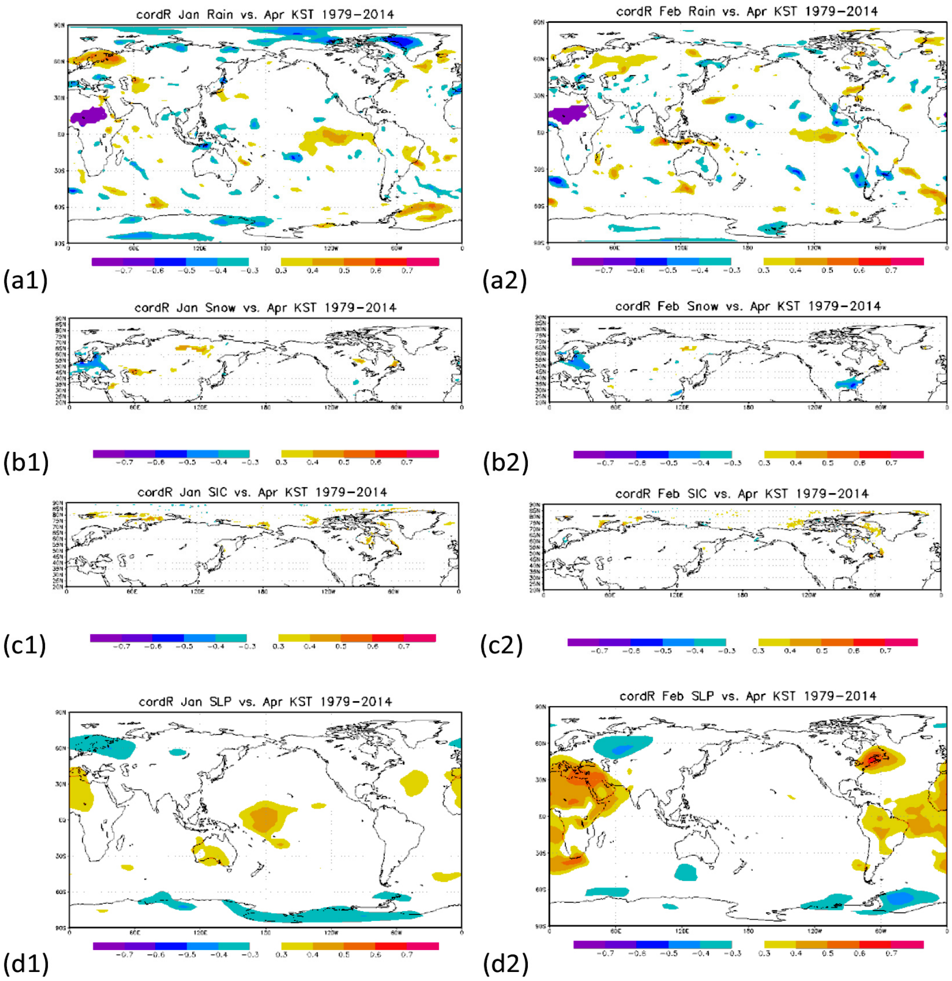


Figure 3.3.2.1 Concurrent correlations between April KST and (a) April precipitation, (b) temperature, (c) sea ice concentration, (d) sea level pressure, (e) sea surface temperature, (f) snow cover, (g) Z300, and (h) Z500, 1979–2014, detrended.

3.3.2.2 Lag Relationships of April Temperature with Preceding January and February Circulation and Surface Properties

Correlation maps between April KST and February and January fields of the variables are shown in Fig. 3.3.2.2. A complex of not-contradicting correlations indicates a shift of the polar front northward, similar to the concurrent correlation maps. The North Atlantic SST pattern resembles that of the positive NAO phase, while circulation correlation patterns indicate an increased extent of the Azore high both westward to America and eastward to the Mediterranean.



(continued on the next page)

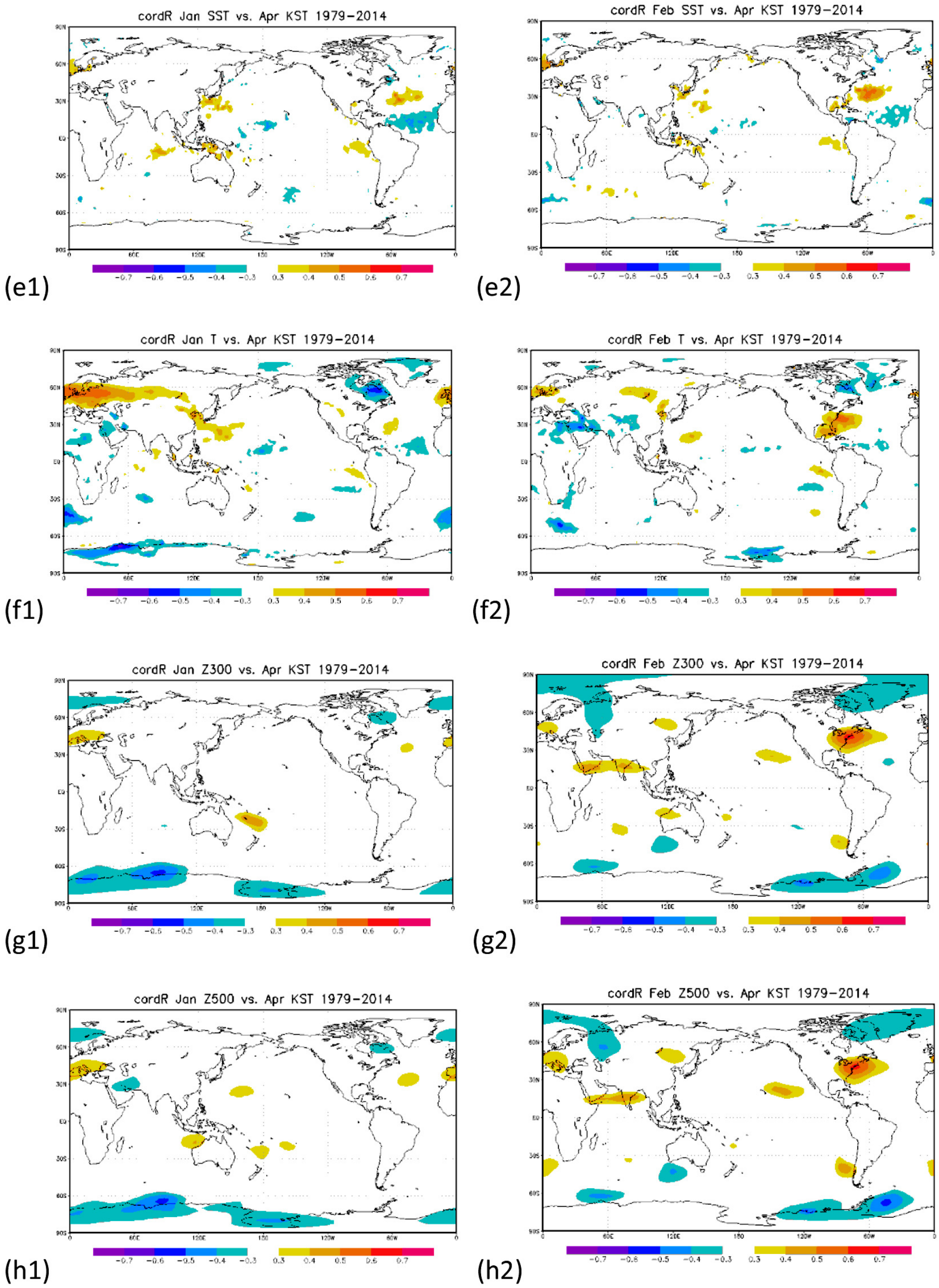


Figure 3.3.2.2 Three- and two-month lag correlations between April KST and (1) January and (2) February for (a) precipitation, (b) temperature, (c) sea ice concentration, (d) sea level pressure, (e) sea surface temperature, (f) snow cover, (g) Z300, and (h) Z500, 1979-2014, detrended.

It is reasonable to construct a prediction scheme based on the February North Atlantic SST or SLP anomalies, which are closely related to each other, since those correspond to global processes; and, accounting for local influences, it is reasonable to use as a second predictor, the January KSST anomaly, with extended coverage of the constructor to make selection more flexible because of the large lead, since correlations are higher and the pattern is wider.

The reasonability of the first predictor (the February North Atlantic SST or SLP anomalies) originates from the close resemblance between the February and April evolution of the upper troposphere zonal wind anomalies associated with the April KST anomalies and February North Atlantic SLP and SST anomalies (Fig. 3.3.2.3). The spatial correlation coefficients between the maps corresponding to each other are shown in Table 3.3.2.1. All three monthly spatial correlations between the maps for the April KST and February SLP exceed 0.7. The SST correlation is less, but not less than 0.42. Mechanisms for the influence of the North Atlantic climate anomalies on the East Asia climate variables have been discussed in a number of studies. Particularly, Watanabe (2004) has shown that the Azore high extending to the middle-eastern Mediterranean in late winter excites a wave train in the upper troposphere. A shift of the polar front jet stream northward associated with the positive polarity of the AO/NAO is obvious. The relationship between the northward shift of the front and the wave train excited by the larger extent of the Azore high and a strong weakening of the East Asia trough is also obvious. It should be emphasized that the February-April evolution of the anomalies of the upper troposphere zonal wind, particularly a northward shift of the polar front jet, is favorable for and indicated by the positive April KST anomaly , which results from and is indicated by the February North Atlantic SST and SLP anomalies presented in Fig. 3.3.2.3.

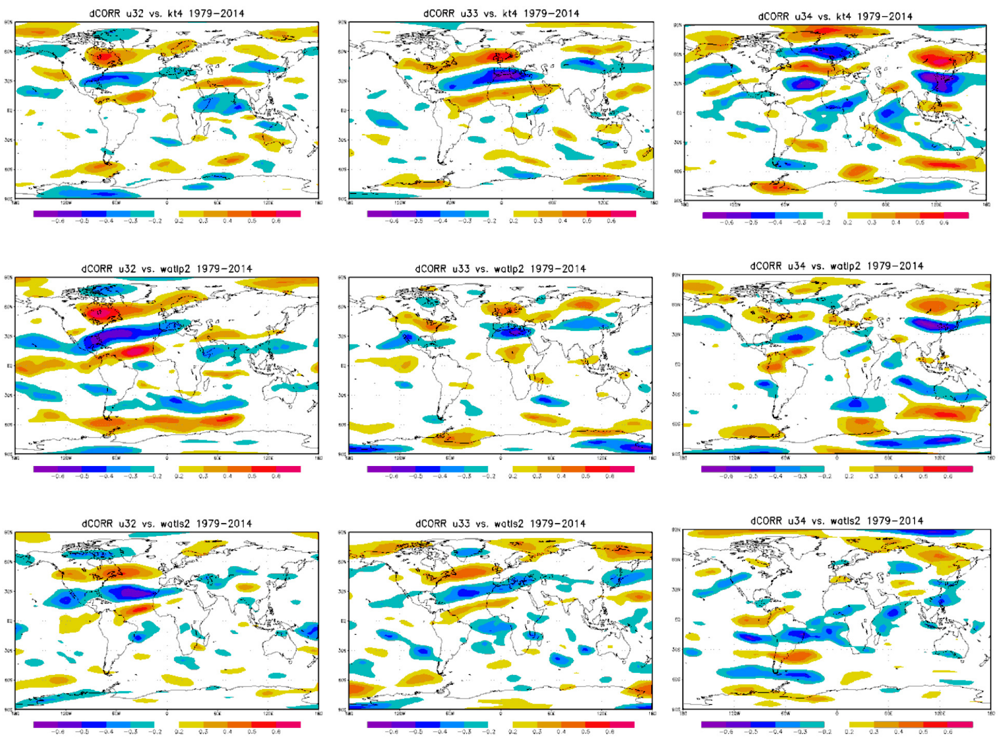
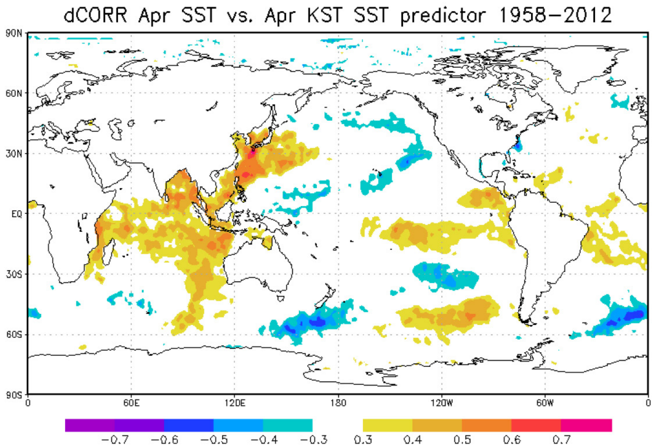


Figure 3.3.2.3 Correlation maps between February (left column), March (central column), and April (right column) monthly U300 fields and April KST (upper row), February SLP averaged over 40–50°N; 80–60°W (central row), and February SST averaged over 25–40°N; 70–40°W (lower row).

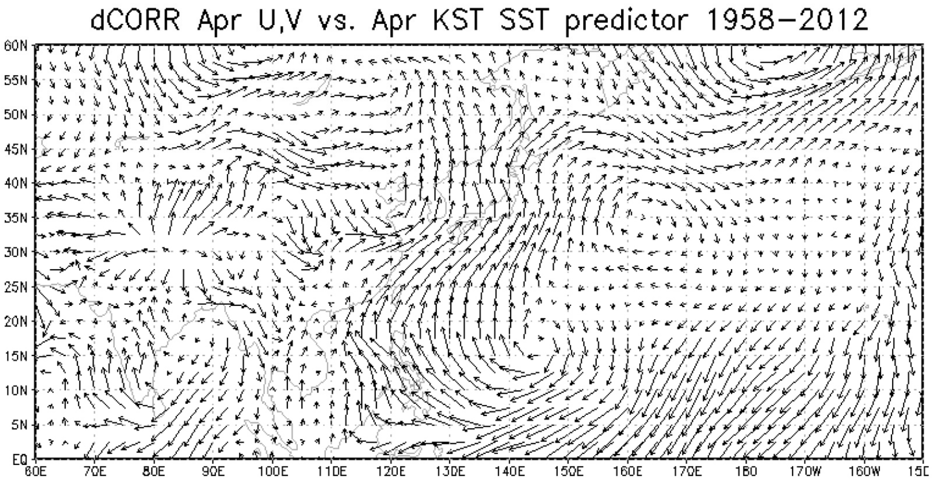
Table 3.3.2.1 Spatial pattern correlation coefficients between the correlation patterns in the maps within the columns shown in Fig. 3.3.2.3, presenting correlations between the grid point U300 monthly series and the series of the predictand, April KST, and the series of North Atlantic February SLP and February SST averaged over 40–50°N; 80–60°W and 25–40°N; 70–40°W, correspondingly. Correlation coefficients estimated over the domain northward of 20°N.

	February U300	March_U300	April_U300
KST _ SLP	0.72	0.71	0.77
KST _ SST	0.50	0.70	0.44
SLP _ SST	0.51	0.42	0.42

The reasonability of the other predictor, the January SST from the western Pacific, is supported by the correlation maps shown in Fig. 3.3.2.4. January predictors for the April KST constructed on the basis of grid point SST series from 45–20°N, 110–140°E leads to April warmer water around Korea, with maximum warming in the East China Sea (Fig. 3.3.2.4a) and southern wind anomalies (Fig. 3.3.2.4b).



(a)



(b)

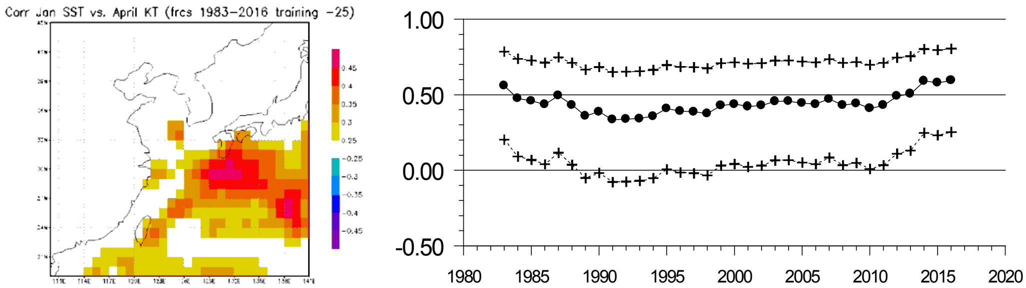
Figure 3.3.2.4 The average of 34 forecast correlation maps of the constructed January SST predictors for April KST with (a) April SST and (b) April wind.

3.3.2.3 Prediction of April Temperature

We have developed two lead time schemes for April. The first is a two-month lead forecast from January. The forecast procedure was similar to that of March. We performed 34 forecasts for the years 1983–2016. The statistical prediction model for all these forecasts is a single predictor regression equation of the form:

$$\text{KST} = k\ \text{KSEP} \qquad \text{(Eq. 3.3.2.3.1)},$$

where KST is the March KST anomaly, KSEP (Korea Seas Extended Predictor) is a predictor constructed from January SST anomalies (Fig. 3.3.2.5a) from the trapezoid 45–20°N, 110–140°E (Korea Seas Extended Anomalies - KSEA); and k is the regression coefficient. This is the basic option, option 0.



(a) (b)

Figure 3.3.2.5 (a) The average of 34 correlation maps used for construction of the predictors. (b) Correlation coefficients (circles) between the predictand and constructed predictors. Correlation coefficient values are plotted in the forecast years, with the correlated series covering the preceding 25-year training periods. Crosses denote the 95% confidence intervals. Real-time forecast scheme.

The series of observed and predicted anomalies of the April KST are shown in Fig. 3.3.2.6. The verification scores of the predictions are shown in Table 3.3.2.1. The forecast verification scheme simulating real-time forecasts gives PCC = 0.47, TCC = 0.50, and MSSS = 0.24. Absolutely independent forecasts are given for 2015 and 2016 (Fig. 3.3.2.6). Therefore, the developed method definitely outperforms the numerical model forecasts. Comparison with other statistical forecasts is impossible because no such verification results have been published in the available journals to date.

Therefore, the developed physical empirical model provides a quite successful two-month lead prediction of April temperature averaged over South Korea.

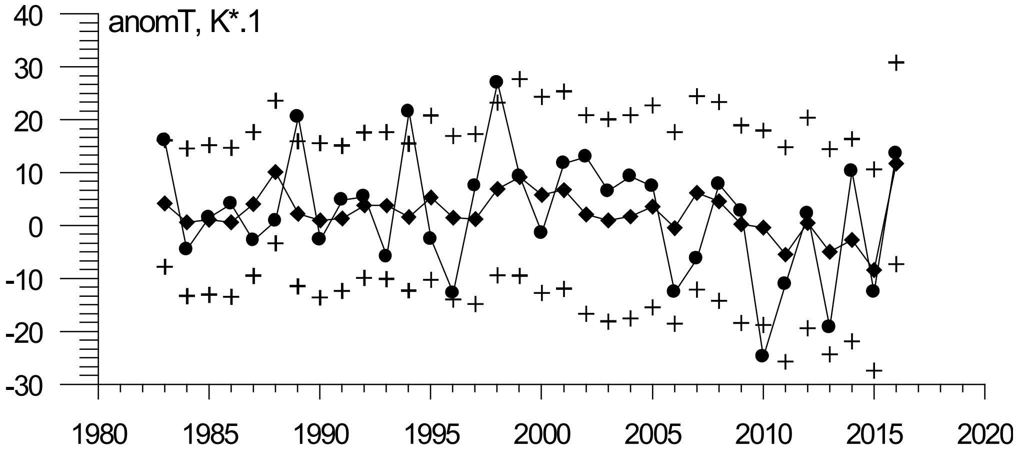


Figure 3.3.2.6 Series of observed (circles) and predicted (rhombs) anomalies of the March KST. Crosses denote the 95% confidence intervals of forecasts. Real-time forecast verification scheme.

Table 3.3.2.1 Skill scores of April forecasts from January based on KSEP.

PCC	TCC	MSSS
0.47	0.50	0.24

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

A possible improvement of the two-month lead forecast is possible from February. For prediction from February, we consider two predictors constructed on the basis of (1) the February North Atlantic SST or SLP grid point series and (2) the January western North Pacific SST. Statistical prediction model for the April KST anomaly forecasts is a multiple regression equation of the form:

$$KST = k1 \ NAPred + k2 \ KSEPed \qquad \text{(Eq. 3.3.2.3.2)}$$

where KST is the April KST anomaly; NAPred is the February North Atlantic predictor, either SST or SLP; KSEP is a predictor constructed from the January KSEA, the same as in Eq. 3.3.2.3.1; and *k1* and *k2* are regression coefficients.

The normalized regression coefficients corresponding to *k1* and *k2* for each forecast are shown in Fig. 3.3.2.7 for (a) SST as the NAPred and for (b) SLP as the NAPred. The average of 34 correlation maps used for construction of the predictors is shown in Fig. 3.3.2.8 for (a) February SST as the NAPred and (b) January KSEP SST.

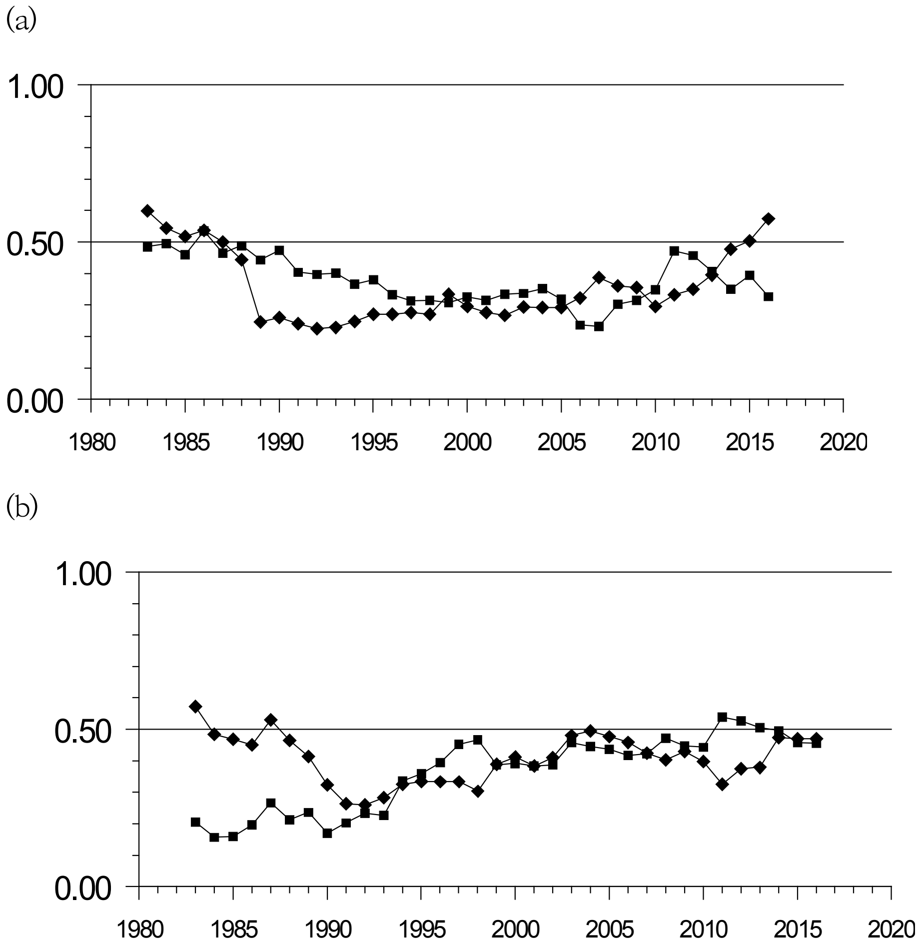


Figure 3.3.2.7 Normalized regression coefficients with (a) SST as the NAPred and (b) SLP as the NAPred: *k1* (NApred) – squares, *k2* (KSST) – rhombuses.

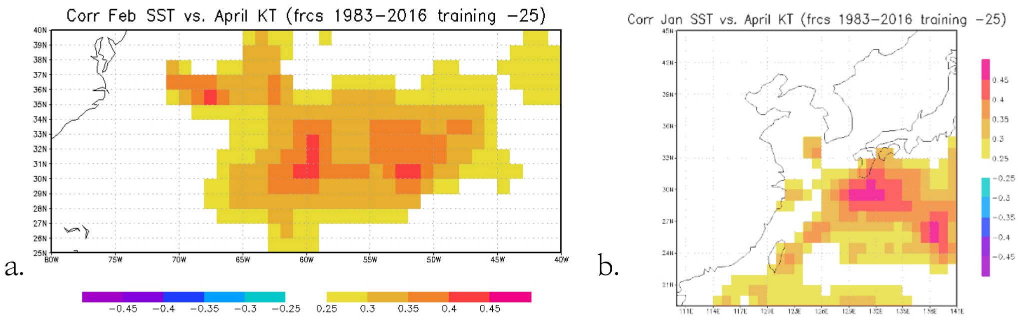


Figure 3.3.2.8 The average of 34 correlation maps used for construction of the predictors based on (a) February North Atlantic SST as the NAPred and (b) January Pacific SST.

The multiple correlation coefficient (square root from determination coefficient) is shown in Fig. 3.3.2.9 for (a) SST as the NAPred and for (b) SLP as the NAPred. These figures demonstrate the difference between the use of SST or SLP as the NAPred in Eq. 3.3.2.3.1. The relationships of the predictor-predictand were rather unstable for the SLP-based predictor, subjectively. They were rather weak for the forecasts before the mid-1990s. However, for the later forecasts the relationships are much stronger, and the multiple correlations for all these forecasts are about 0.6. Meanwhile, the SST based predictor relationships are quite stable, although not so strong as for SLP after the mid-1990s.

The use of SLP instead of SST essentially improves the skill of the forecasts. As shown in Tables 3.3.2.2 and 3.3.2.3, the use of SLP increases all the verification scores by 0.1–0.2. Which of these predictors is preferable? There is no a single answer to this question. The use of the SST-based predictor deteriorates the skill but makes the extrapolation of the skill to the (nearest) future more reliable. The use of the SLP-based predictor essentially improves the skill because of the very strong relationships during the recent decades and the very skillful forecasts of the recent years; however, if the relationships return to the level of the 1950s-1960s, deterioration may appear. Therefore, this prediction model should be used carefully with a check of the current strength of the relationships. However, it should be noted that strong relationships during the training period do not guarantee a skillful real-time forecast.

Table 3.3.2.2 Skill scores of April forecasts with SST as the NAPred.

PCC	TCC	MSSS
0.51	0.51	0.24

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

Table 3.3.2.3 Skill scores of April forecasts with SLP as the NAPred.

PCC	TCC	MSSS
0.63	0.65	0.40

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

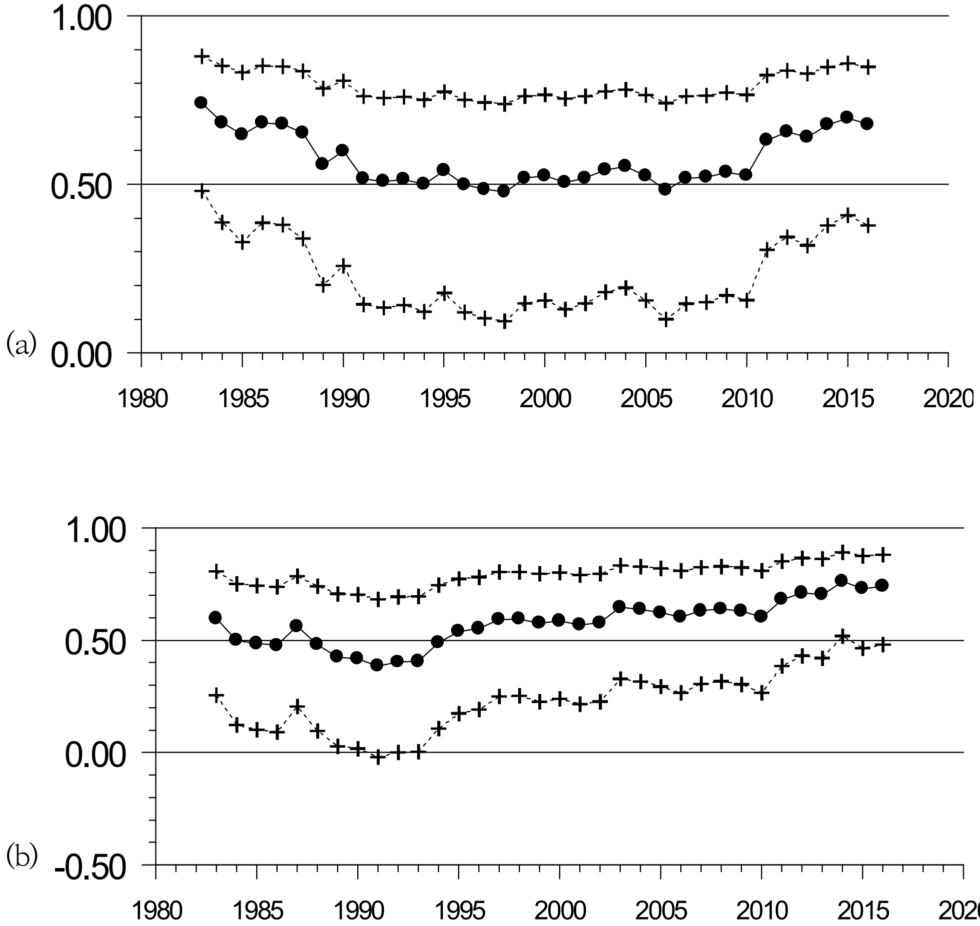
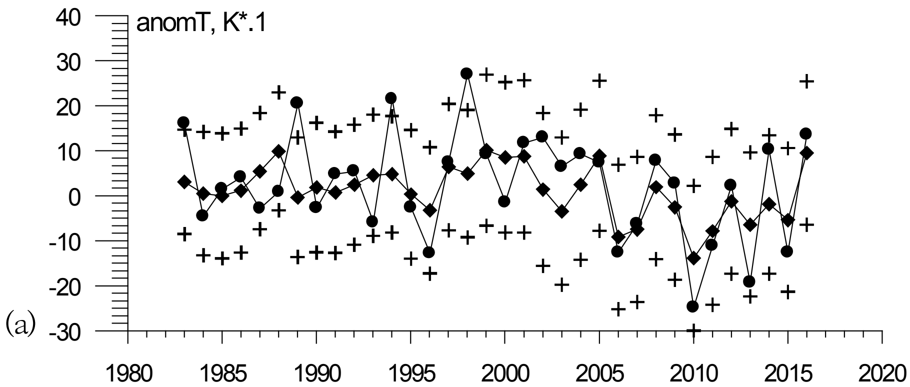


Figure 3.3.2.9 Multiple correlation coefficients (circles) corresponding to Eq. 3.3.2.3.1 with (a) SST as the NAPred and (b) SLP as the NAPred. Crosses are the 95% confidence intervals.

The observed and predicted series of the April KST anomaly are shown in Fig. 3.3.2.10. Both NAPreds work properly, especially during recent years, including the years of 2015 and 2016, i.e., the years with an absolutely independent forecast.



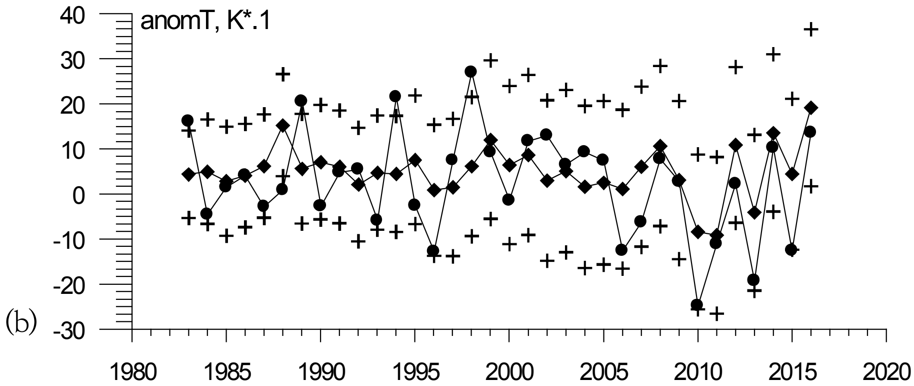
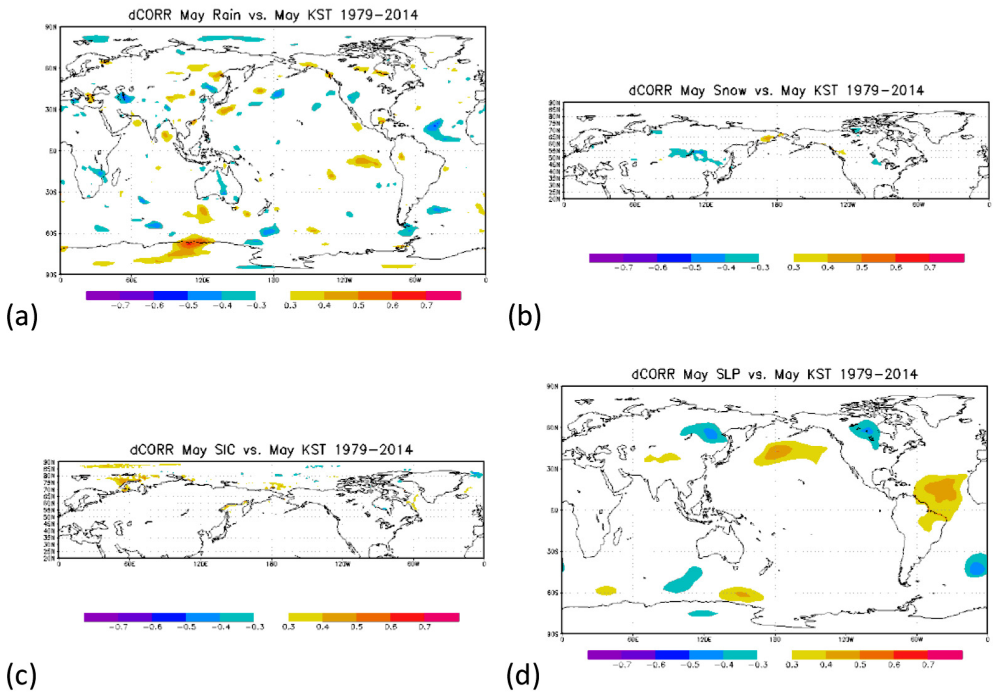


Figure 3.3.2.10 Series of observed (circles) and predicted (rhombs) anomalies of April KST. Crosses denote the 95% confidence intervals of the forecasts. Real-time forecast verification scheme. (a) SST as the NAPred and (b) SLP as the NAPred.

3.3.3. May

3.3.3.1 Concurrent Relationships of May Temperature with Circulation and Surface Properties

As compared to March and April, the correlations of May temperature in South Korea with circulation variables are rather weak. May temperature anomalies slightly persist from April (correlation of 0.44 and grid point correlations of up to 0.5 in the southwestern provinces) and are related to the general circulation, which may be illustrated by the consistent correlation patterns of SLP, precipitation, snow, and temperature (Fig. 3.3.3.1), suggesting an intensification and possible northward shift of the polar front. Conversely, with a midday May sun angle essentially exceeding 60°, the May anomalies should be related to anomalies of insolation and controlled by cloudiness. Such a double influence is also supported by an insignificant correlation between May temperature and precipitation (-0.14); although it is negative, indicating a summer-like relationship. A stable air stratification, that is, a downward air movement associated with an anticyclonic circulation, favors enhanced insolation. The positive correlation patterns over Korea at Z500 and Z300 indicate positive geopotential height anomalies, which maintain downward movement, but result from the heating below.



(continued on the next page)

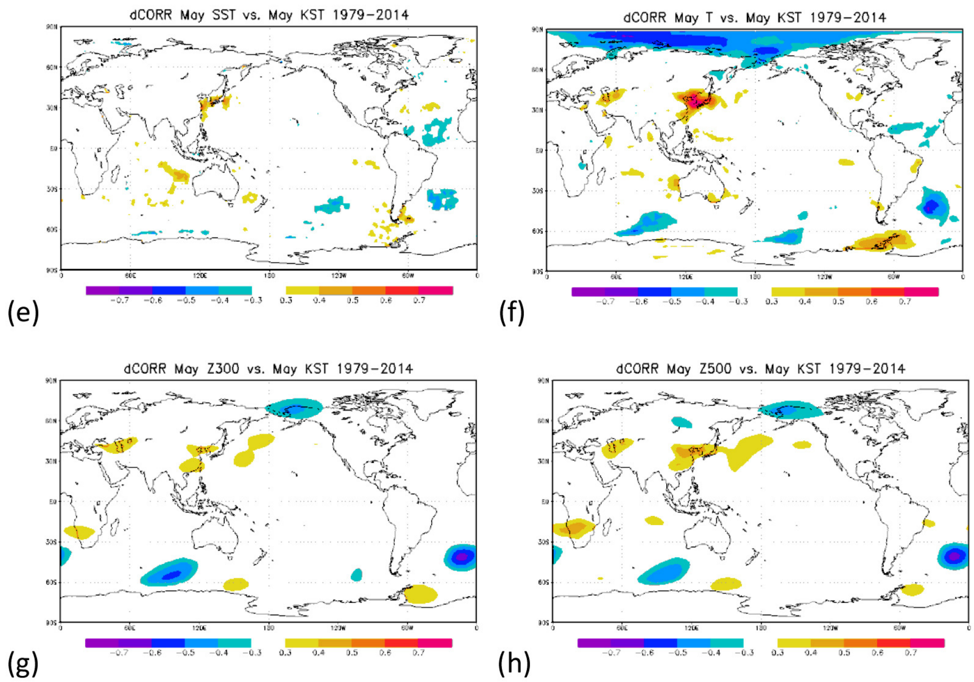


Fig. 3.3.3.1 Concurrent correlations between May KST and (a) May precipitation, (B) temperature, (c) sea ice concentration, (d) sea level pressure, (e) sea surface temperature, (f) snow cover, (g) Z300, and (h) Z500, 1979-2014, detrended.

3.3.3.2 Lag Relationships of May Temperature with Preceding March Circulation and Surface Properties

The two-month lag correlation maps corresponding to the one-month lead forecasts (Fig. 3.3.3.2) between May KST and March circulation and surface variables demonstrate very low significance for the relationships as compared with the previous spring months. In addition, the relationships between May KST and climate variables of the months preceding March also demonstrate low significance. Nonetheless, correlation patterns are recognizable, not high but significant, between the May KST and March precipitation, SST, and temperature. These resemble the inverted patterns of negative SST anomalies and positive SLP anomalies over the Philippine Sea, which are associated with El-Nino and persist from the autumn through early summer as discussed in a number of studies, e.g., Wang et al. 2000; Lau et al., 2005; and Yuan et al., 2012.

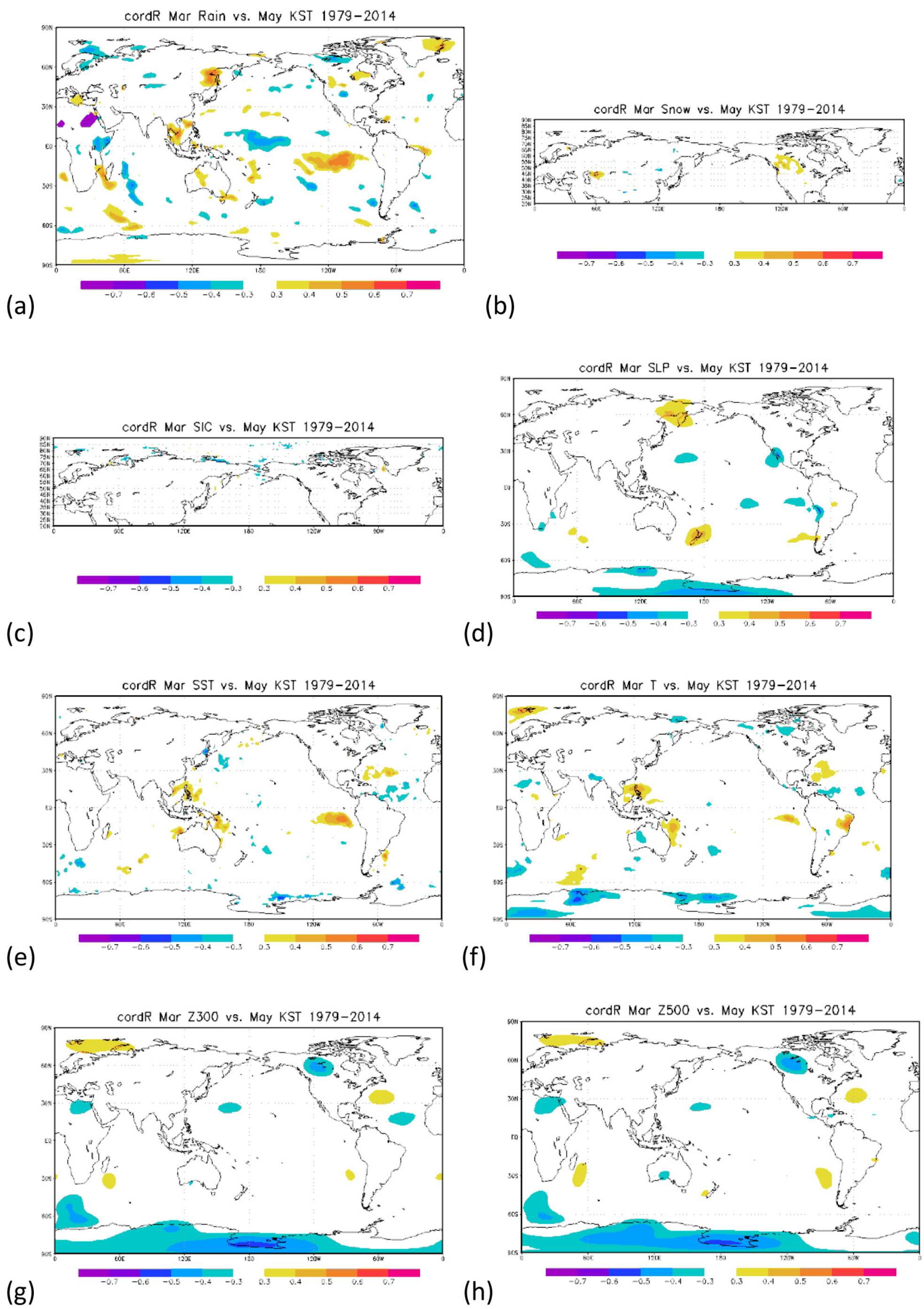


Fig. 3.3.3.2 Two-month lag correlations between May KST and (a) March precipitation, (b) temperature, (c) sea ice concentration, (d) sea level pressure, (e) sea surface temperature, (f) snow cover, (g) Z300, and (h) Z500, 1979-2014, detrended.

Since the angle of the Sun exceeds 70° over Korea in May, from general considerations, it follows that the May temperature anomalies are defined by anomalies of solar irradiance rather than anomalies in heat advection. The solar irradiance anomalies are controlled by cloudiness. In addition, late spring and summer rains lead to a temporal temperature decrease due to the heat absorption by evaporating water. Therefore, from these considerations, it follows that positive (negative) temperature anomalies favor anticyclonic (cyclonic) circulation providing divergence (convergence) in the lower troposphere. This is in accord with the circulation anomaly of the inverted sign over the Philippine Sea. Therefore, it is reasonable to use the Philippine Sea March SST anomaly as a predictor constructor for May KST. A map of the average correlation of 34 forecasts between the March predictors for May KST and May wind components is shown in Fig. 3.3.3.3. It demonstrates an anticyclonic circulation pattern over Korea.

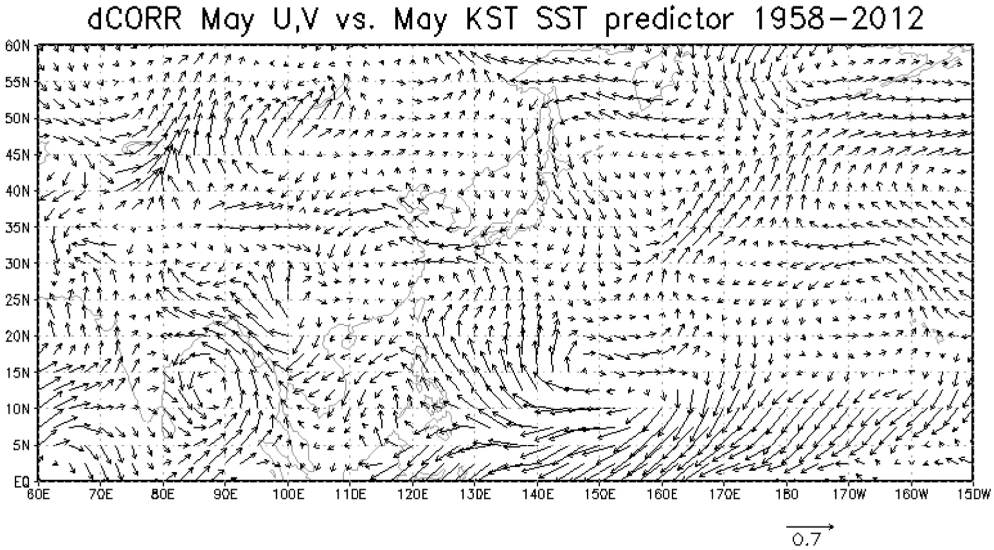


Figure 3.3.3.3 Map of correlations averaged over 34 forecasts between the constructed March SST predictors for May KST and May wind.

3.3.3.3 Prediction of May Temperature

For the prediction of May temperature from January, we have extended the SST constructor area to cover the whole western North Pacific, the trapezoid 45–0°N, 110–140°E. Instead of coming from the seas surrounding Korea, the main, although quite weak, impact comes from the Philippine Sea area. This indicates that the influence of the winter conditions on the following spring temperature in Korea decays by May, and the May temperature variability is dominated by other causes. This would be expected based on studies of the persistent ENSO related anomaly of SST in the Philippine Sea (e.g., Wang et al. 2000; Lau et al., 2005; Yuan et al., 2012). The correlation coefficients between the predictand and constructed predictors are shown in Fig. 3.3.3.4. The prediction of May temperature from January is a three-month lead forecast. It is not too surprising that the skill of the forecasts is low (Figure 3.3.3.5 and Table 3.3.3.1).

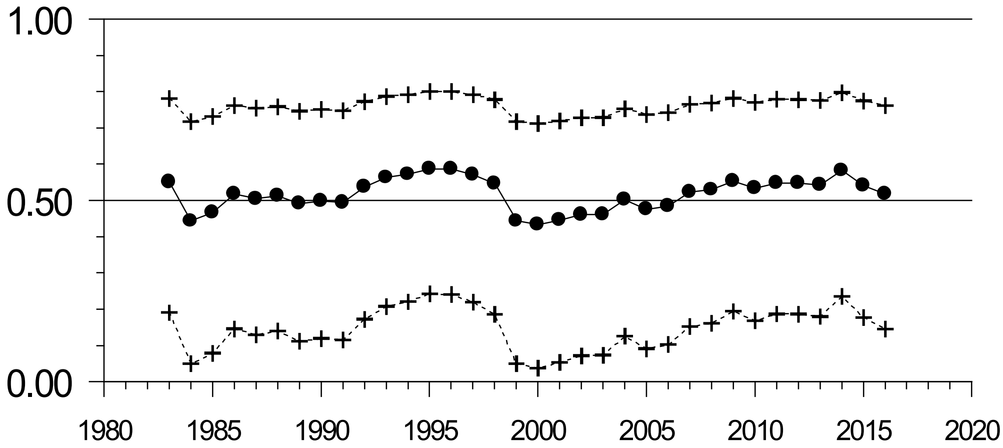


Figure 3.3.3.4 Correlation coefficients (circles) between the predictand and constructed predictors. Correlation coefficient values are plotted in the forecast years, with correlated series covering the preceding 25-year training periods. Crosses denote the 95% confidence intervals. Real-time forecast scheme.

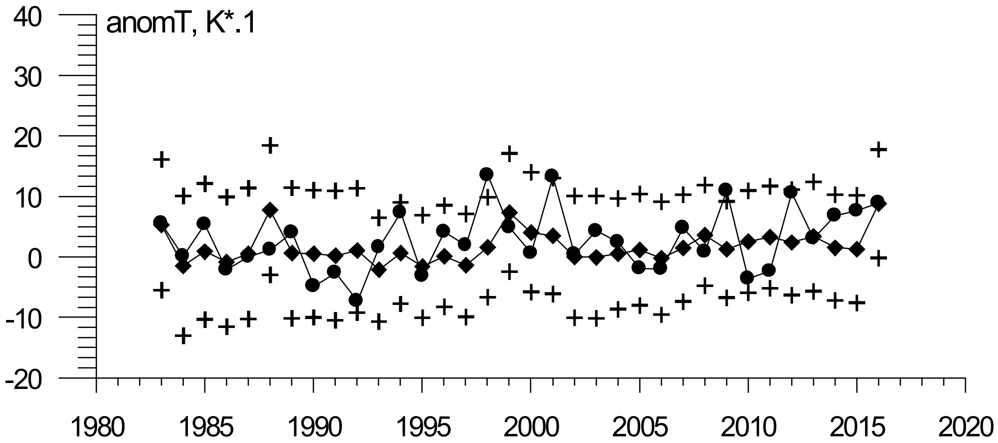


Figure 3.3.3.5 Series of observed (circles) and predicted (rhombs) anomalies of the May KST from January. Crosses denote the 95% confidence intervals of forecasts. Real-time forecast verification scheme.

Table 3.3.3.1 Skill scores of May forecasts based on January SST.

PCC	TCC	MSSS
0.17	0.42	0.14

Since the Philippine Sea anomaly persists from autumn through early summer (Wang et al. 2000; Lau et al., 2005; Yuan et al., 2012), we use the Philippine Sea March SST anomaly as a predictor constructor in a single predictor regression equation for the prediction of May KST from March. The average of 34 correlation maps used for the construction of the predictors is shown in Fig. 3.3.3.6.a, and the normalized regression coefficients associated with each of the 34 forecasts are shown in Fig. 3.3.3.6b.

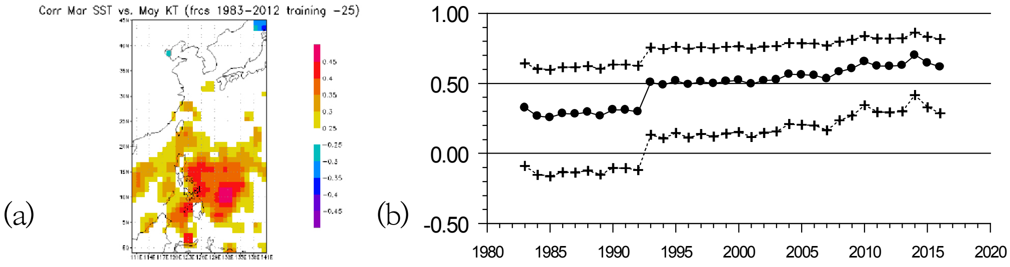


Figure 3.3.3.6 (a) The average of 34 correlation maps used for the construction of the predictors. (b) Correlation coefficients (circles) between the predictand and constructed predictors. Correlation coefficient values are plotted in the forecast years, with correlated series covering the preceding 25-year training periods. Crosses denote the 95% confidence intervals. Real-time forecast scheme.

The series of observed and predicted May KST are shown in Fig. 3.3.3.7, and the skill scores of the predictions of May KST are shown in Table 3.3.3.2. They are lower than those for March and April, which is expected because of the absence of strong predictors. Nonetheless, the score metric values are statistically significant at the 95% confidence level at least.

Table 3.3.3.1 Skill scores of May forecasts based on the March SST.

	PCC	TCC	MSSS
Forecast simulation	0.34	0.52	0.28

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

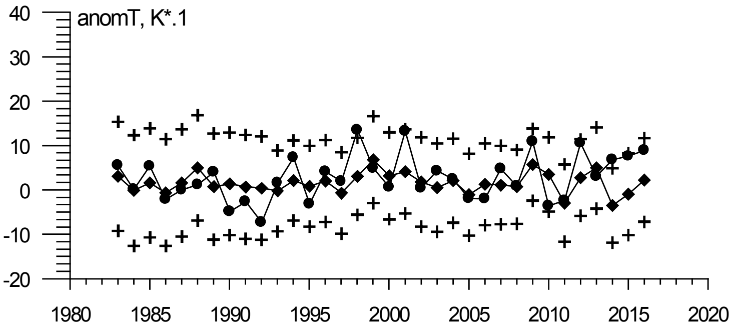


Figure 3.3.3.7 Series of observed (circles) and predicted (rhombs) anomalies of the May KST from March. Crosses denote the 95% confidence intervals of the forecasts. Real-time forecast verification scheme.

3.3.4 March-May

As mentioned in Section 3.3.1, the seasonal KST anomaly must to a large degree be defined by the March anomaly so that prediction of the seasonal mean anomaly is possible based on predictors for the March anomaly. The constructed predictors are close to those obtained for the prediction of the March KST anomaly. Although the level of correlations (Fig. 3.3.4.1) is slightly lower, the verification scores are on a similar level, slightly outperforming in real-time forecasts and slightly underperforming in cross validated forecasts (Table 3.3.4.1). The series of observed and predicted MAM anomalies are shown in Fig. 3.3.4.2. The predictions for 2015 and 2016 are absolutely independent.

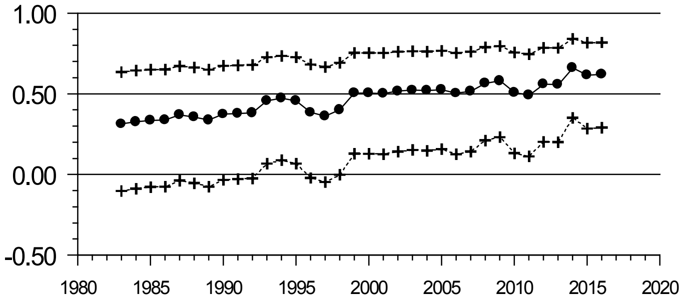


Figure 3.3.4.1 Correlation coefficients (circles) between the predictand and constructed predictors. Correlation coefficient values are plotted in the forecast years, with correlated series covering the preceding 25-year training periods. Crosses denote the 95% confidence intervals. Real-time forecast scheme.

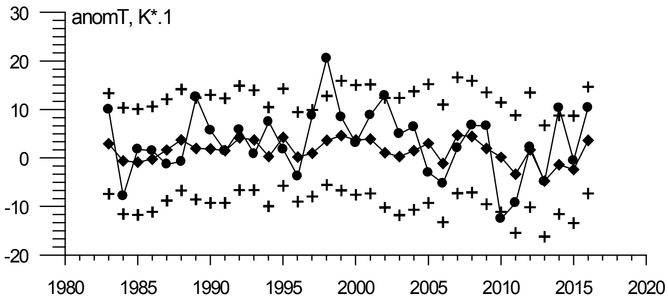


Figure 3.3.4.2 Series of observed (circles) and predicted (rhombs) anomalies of the MAM KST Crosses denote the 95% confidence intervals of forecasts. Real-time forecast verification scheme.

Table 3.3.4.1 Performance characteristics for MAM forecasts.

PCC	TCC	MSSS
0.46	0.58	0.29

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

3.4 Probabilistic Interpretation

We have developed probabilistic forecasts on the basis of deterministic forecast accounting for the forecast uncertainty. The estimation of deterministic forecast uncertainty is detailed by e.g., Taylor (1982). Its probabilistic interpretation in the context of long range forecasting is detailed by Min et al. (2011).

The results of the verification of the probabilistic forecasts, developed on the basis of the deterministic real-time forecasts, are shown in Table 3.4.1. The abbreviations not yet given are those commonly accepted: BN, NN, and AN for below, near, and above normal.

Table 3.4.1 Skill scores of probabilistic forecasts.

	RPSS	ROCS BN	ROCS NN	ROCS AN
March (from Jan.)	0.28	0.69	0.59	0.84
April (from Jan.)	0.17	0.75	0.51	0.63
April (from Mar.: NASST&KSEP)	0.16	0.80	0.36	0.62
April (from Mar.: NASLP&KSEP)	0.30	0.90	0.56	0.75
May (from Jan.)	0.07	0.45	0.52	0.57
May(from Mar.)	0.15	0.62	0.62	0.69
MAM (from Jan.)	0.26	0.92	0.42	0.68

The RPSS values are positive, indicating that the developed forecasts are more skillful than the climatological forecasts. The ROCS values show that the skill of the forecast probabilities in the AN and BN categories is quite good. The performance of the forecasts of probabilities in the NN category are not so skillful, which is typical and explained by van der Dool and Toth (1991).

These probabilistic forecasts from the developed physical-empirical models can be easily combined with APCC probabilistic forecasts from dynamic models. The individual model weight in APCC probabilistic MME forecast is inversely proportional to the random error in the ensemble mean. Similar weight for the statistical forecast could be the weight inversely proportional to the square root of the training period sample size.

4. CONCLUDING REMARKS

This study has focused on the development of physical-empirical models (statistical models based on physically plausible relationships) for the prediction of spring seasonal mean and monthly temperatures over East Asia, particularly South Korea. These models are appropriate to use with APCC operational practice as a support and to complement APCC MME forecasts.

The statistical tool underlying the physical-empirical models is regression. The prediction scheme is unique because of the use of the most appropriate predictors for performing the forecast. This is achieved by flexible construction of the predictors in accordance with peculiarities of the given training period. Predictors are constructed for each forecast independent of previous forecasts.

The obtained skill for South Korea March and April temperatures is provided mainly by the extratropical predictors related to the EAWM originating from the global winter processes. Therefore, it is reasonable to combine forecasts from the developed physical-empirical models with forecasts from the dynamic models, which are mainly governed by the tropical SST anomalies. Although May temperature prediction is based mainly on the processes originating in the tropics, similar to the dynamic models, it could also be used to complement the APCC MME forecasts.

REFERENCE

Aristov, N.A., N.A. Bagrov, K.A. Vasyukov, G.G. Gromova, N.I.Zverev, A.L. Kats, S.T.Pagava (editor), 1972: Guidebook on Monthly Weather Forecasts (in Russian) . Gidrometeoizdat, Leningrad, 366 p.

Bagrov, N.A., 1959: Analitical presentation of a succession of meteorological fields by empirical (natural) orthogonal components, *Tr. Centr. Inst. Of Forecasting.,* **74,** 3-24.

________, 1969: On the equivalent number of independent data (in Russian). *Tr. Gidrometeor. Cent.,* **44,** 3-11.

_______, K.V. Kondratovich, D.A. Hed, A.I. Ugryumov, 1985: Long range meteorological forecasts (in Russian). Gidrometeoizdat, Leningrad, 248 p.

Bretherton C, Widmann M, Dymnikov V, Wallace J, Blade I. 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Clim.* **12**: 1990-2009.

Clark, M.P., M.P. Serreze, D.A., Robinson, 1999: Atmospheric controls on Eurasian snow extent. International Journal of Climatology, 19, 27-40.

Compo, G.P., J.S. Whitaker, P.D. Sardeshmukh, and coauthors, 2011: The Twentieth Century Reanalysis Project. Quarterly Journal of Royal Meteorological Society, **137:** 1-28. DOI: 10.1002/qj.776.

________, J.S. Whitaker, and P.D. Sardeshmukh,. 2006: Feasibility of a 100 year reanalysis using only surface pressure data. Bulletin of American Meteorological Society, **87**: 175-190.

Doblas-Reyes, F.J., 2010: Seasonal prediction over Europe, ECMWF Seminar on Predictability in the European and Atlantic regions, 6 to 9 September 2010, 171-185

Girs, A.A., 1960: The bases of long range weather forecasts (in Russian). Gidrometeoizdat, Leningrad, 570 p.

_______, 1971: Multiyear variations of atmospheric circulation and long range hydrometeorological forecasts (in Russian). Gidrometeoizdat, Leningrad, 280 p.

Groisman, P.Y., T.E. Karl, R.W. Knight, G.L. Stenchikov, 1994: Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere. - J. Clim., 7, 1633-1656

Hall R, Erdélyi R, Hanna E, Jones JM, Scaife AA. 2015. Drivers of North Atlantic polar front jet stream variability. *Int. J. Climatol.* **35**: 1697-1720, doi: 10.1002/joc.4121

Harris, I., P.D. Jones, T.J. Osborn, and D.H. Lister, 2014: Updated high-resolution grids of monthly climatic observations. In press, Int. J. Climatol. **34**: 623-642 Doi: 10.1002/joc.3711

Huang, D.-Q., J. Zhu, Y.-C. Zhang, and A.-N. Huang, 2014: The Different Configurations of the East Asian Polar Front Jet and Subtropical Jet and the Associated Rainfall Anomalies over Eastern China in Summer. J. Clim. 27, 8205 – 8220.

Jhun, J.G., and E.J. Lee, 2004: A new East Asian winter monsoon index and associated characteristics of the winter monsoon. Journal of Climate **17**: 711-726.

Kalnay, E. and coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. Bulletin of American Meteorological Society **77:** 437- 471.

Khromov, S.P., 1985: Meteorology and climatology for geographical schools, Gidrometeoizdat, Leningrad, 455 pp.

Kim, G., J.B. Ahn, V.N. Kryjov, S.J. Sohn, W.T. Yun, R. Graham, R.K. Kolli, A. Kumar, and J.P. Ceron, 2016: Global and regional skill of the seasonal predictions by WMO Lead Centre for Long-Range Forecast Multi-Model Ensemble. - International Journal of Climatology, V36, 1657-1675.

Kryjov, V.N., 2002. The influence of the winter Arctic oscillation on the northern Russia spring temperature. *Int. J. Climatol.* **22**: 779-785, doi: 10.1002/joc.746

_______, 2004: Searching for circulation patterns affecting Northern Europe annual temperature. Atmos. Sci. Lett. **5**: 23-34, doi: 10.1016/ j.atmoscilet.2003.11.00.

_______, 2015: Development of physical-empirical models for seasonal predictions of wintertime climate variables over East Asia, APCC Research Report-2015.

_______, and Y.-M. Min 2016: Predictability of the wintertime Arctic Oscillation based on autumn circulation. - International Journal of Climatology, V36, 4181-4186.

Kryzhov, V.N., 2012: Downscaling of the global seasonal forecasts of Hydrometcenter of Russia for North Eurasia. Russian Meteorology and Hydrology, **37**: 291-297

Kug, J.S., J.Y. Lee, and I.S. Kang, 2008: Systematic error correction of dynamical seasonal prediction of sea surface temperature using a stepwise pattern projection method. Monthly Weather Review **136**: 3501-3512, doi: 10.1175/2008MWR2272.1.

Lau, N.C., Leetmaa A., Nath M.J., Wang H.L., 2005: Influences of ENSO-induced Indo-western Pacific SST anomalies on extratropical atmospheric variability during the boreal summer. J Clim 18:2922-2942

Lee, J.Y., S.S. Lee, B. Wang, K.J. Ha, and J.G. Jhun, 2013: Seasonal prediction and predictability of the Asian winter temperature variability. Climate Dynamics **41**, 573-587, DOI 10.1007/s00382-012-1588-5.

Luo, X., and Y. Zhang, 2015: The Linkage between Upper-Level Jet Streams over East Asia and East Asian Winter Monsoon Variability. J. Clim., 28, 9013-9028.

Li, L., and Y. Zhang, 2014: Effects of Different Configurations of the East Asian Subtropical and Polar Front Jets on Precipitation during the Mei-Yu Season. J. Clim., 27, 6660 - 6672.

Min, Y.-M., V.N. Kryjov, and J.-H. Oh, 2011: Probabilistic Interpretation of Regression-Based Downscaled Seasonal Ensemble Predictions with the Estimation of Uncertainty. - J. Geophys. Res., 116, D08101, doi:10.1029/2010JD015284

________, _______, and S.M. Oh, 2014: Assessment of APCC Multi-Model Ensemble Prediction in Seasonal Climate Forecasting: Retrospective (1983-2003) and Real-time Forecasts (2008-2013) Journal of Geophysical Research-Atmospheres, 119(21):12,132 -12,150. DOI:10.1002/2014JD022230

Murphy, A.H., 1988: Skill scores based on the mean squared error and their relationships to the correlation coefficient. Monthly Weather Review, 116: 2417-2424.

Rayner, N.A., D.E. Parker, E.B. Horton, and coauthors, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century Journal of Geophysical Research **108**: D14, 4407 10.1029/2002JD002670.

Rigor IG, Wallace JM, Colony RL. 2002. On the response of sea ice to the Arctic Oscillation. *J. Clim.* **15**: 2648-2663, doi: 10.1175/1520-0442(2002)015<2648:ROSITT>2.0.CO;2

Taylor, J. R. (1982), An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, 327 pp., Univ. Sci., Sausalito, Calif.

Ugryumov, A.I., 1981: Heat regime of ocean and long range weather forecasts (in Russian). Gidrometeoizdat, Leningrad, 175 p.

van den Dool, H., and Z. Toth (1991), Why do forecasts for “near normal” often fail?, Weather Forecast., 6, 76-85, doi:10.1175/1520-0434(1991) 006<0076:WDFFNO> 2.0.CO;2.

Von Storch, H. and F.W. Zwiers (1999) Statistical Analysis in Climate Research. Cambridge University Press, Cambridge.

Wang B, Wu R, and Fu X (2000) Pacific-East Asia teleconnection: how does ENSO affect East Asian climate? J Clim 13:1517-1536

Wang B, Xiang B, and Lee JY (2013) Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. PNAS 10:2718–2722

Watanabe, M., 2004: Asian jet waveguide and a downstream extension of the North Atlantic Oscillation. J Clim 17,4674–4691.

Whitaker, J.S., G.P. Compo, X. Wei, and T.M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. Monthly Weather Review, **132**: 1190–1200.

Wilks DS. 1995. *Statistical Methods in the Atmospheric Sciences*, 2ndedn.ElsevierAcademicPress: SanDiego,CA

WMO, 2012: Manual on the Global Data-processing and Forecasting System. Volume I – Global Aspects, Updated in 2012. WMO–No. 485, Geneva

Wu, R., and J. L. Kinter III, 2010: Atmosphere-ocean relationship in the midlatitude North Pacific: Seasonal dependence and east-west contrast, J. Geophys. Res., 115, D06101, doi:10.1029/2009JD012579.

______, G. Liu, and Z. Ping (2014), Contrasting Eurasian spring and summer climate anomalies associated with western and eastern Eurasian spring snow cover changes, J. Geophys. Res. Atmos., 119, 7410–7424, doi:10.1002/2014JD021764.

______, and B. P. Kirtman, 2007: Observed Relationship of Spring and Summer East Asian Rainfall with Winter and Spring Eurasian Snow, . J. Clim., 20, 1285 – 1304

Wu Z, Wang B, Li J, Jin FF, 2009: An empirical seasonal prediction of the East Asian summer monsoon using ENSO and NAO. J Geophys Res 114:D18120. doi:10.1029/2009JD011733

Yim S.-Y., Wang B., Kwon M. (2013) Interdecadal change of the controlling mechanisms for East Asian early summer rainfall variation around the mid-1990s. Clim Dyn 42:1325–1333

______, B. Wang, W. Xing, and M.-M. Lu, 2015: Prediction of Meiyu rainfall in Taiwan by multi-lead physical-empirical models. Clim Dyn., 44:3033-3042

Ye, H.,, and Z. Bao, 2001: Lagged teleconnections between snow depth in northern Eurasia, rainfall in southeast Asia and sea-surface temperatures over the tropical pacific ocean, I*nt. J. Climatol*. **21**: 1607–1621, doi: 10.1002/joc.695

Yuan, Y., S. Yang, Z. Zhang, 2012: Different Evolutions of the Philippine Sea Anticyclone between the Eastern and Central Pacific El Nin˜ o: Possible Effects of Indian Ocean SST. J. Clim, 25, 7867 – 7883,

Yudin, M.I., A.V. Mesherskaya, 1977: Complex physical-statistical method of prediction with long lead time, Meteorologia I Gidrologia (Meteorology and Hydrology, Issue 1, 3–12.

APPENDIX 1

PREDICTION METHOD

Prediction of a future series value is performed by means of multiple regression:

$$y_t = b_0 + \sum_{k=1}^{K} b_k x_{k,t} + \varepsilon_t \tag{A1.1}$$

where y_t is the predictand at time t, $x_{k,t}$ is the k-predictor at time t, and ε_t is the error at time t. Accounting for the size of a real available training series of 25 - 30 years, the number of predictors, K, should not exceed two.

The predictand is a target variable; in our case, it is the temperature averaged over South Korea.

Predictor(s) are estimated based on the training period relationships by projecting the predictor variable fields onto a correlation map between the predictand and predictor variable fields. This method, with some variations, has been being utilized in statistical forecast models for decades. Currently, it is widely applied in downscaling procedures (e.g., Kug et al., 2008; Min et al., 2011; Kryzhov, 2012). This method is applied by Lee et al. (2013) in "Predictable Modes" statistical forecasting technology, which marked a recovery of the physical-empirical approach to long range forecasting. The main difference between the developed method and the methods mention above is the absence of any restrictions on the correlation map values, which provides higher flexibility in the construction of the predictors.

The prediction procedure consists of five steps.

Step 1. We compute a correlation map of October predictor fields z and the DJF target variable y series using the data from the training period of size T years:

$$r_{k,s} = \frac{\sum_{t=1}^{T} z_{k,s,t} y_t}{(\sum_{t=1}^{T} z_{k,s,t}^2 \sum_{t=1}^{T} y_t^2)^{0.5}}, \tag{A1.2}$$

where s is a space (grid point) counter.

Step 2. We construct a T-year long series of predictor(s), computed as a projection of the October predictor(s) anomaly fields on the correlation map.

$$x_{k,t} = \sum_{s=1}^{S} r_{k,s} z_{k,s,t} \tag{A1.3}$$

If we use multiple regression, i.e., a set of predictors, the first and the second steps are performed for each k-predictor separately.

Step 3. We derive a regression equation by means of a least squares estimation (minimizing the mean squared error ε) based on the T-years training period.

Step 4. We estimate the $(T+1)^{th}$ year predictor(s) $x_{k,T+1}$ by projecting the October predictor variable(s) field(s) from the $(T+1)^{th}$ year onto the (corresponding) correlation map(s).

$$x_{k,T+1} = \sum_{s=1}^{S} r_{k,s} z_{k,s,T+1} \tag{A1.4}$$

Step 5. We estimate the forecast value of the DJF target variable using the $(T+1)^{th}$ year value(s) of the predictor(s).

$$y_{T+1} = b_0 + \sum_{k=1}^{K} b_k x_{k,T+1} \tag{A1.5}$$

Unique to this method is that there are no restrictions on the correlation values. In such an approach, a predictor time series becomes collinear to the expansion coefficient series obtained by singular value decomposition of a correlation matrix equal to the corresponding correlation map. This makes the method flexible.

APPENDIX 2

TREND RELATED ISSUES

The original time series of the South Korean temperature contains a trend component (Figs. 3.1.1 and 3.1.2), giving rise to a number of questions. Can the developed forecast model outperform simple forecasts by extrapolation into the forecast year of the trend from the training period? What is the contribution of the short term variability component to the performance of the forecasts and what is the contribution of the trend component? Can the short term component be predictable if the trend does not affect derivation of the regression equations, i.e., it does not exist in all the training periods plus one year? Is there predictability for the high-pass filtered series?

For convenience of comparison, the verification scores of the forecasts obtained for the original series with climatological forecasts as reference are summarized in Table A2.1.

Table A2.1 Verification scores of the developed forecasts.

	PCC	TCC	MSSS
March			
Forecast simulation	0.45	0.56	0.25
Cross-Validation (-9 yrs.)	0.65	0.65	0.42
April (AtlSST)			
Forecast simulation	0.51	0.51	0.24
Cross-Validation (-9 yrs.)	0.47	0.47	0.17
May			
Forecast simulation	0.34	0.52	0.28
Cross-Validation (-9 yrs.)	0.62	0.62	0.38

1. Can the developed forecast model outperform trivial forecasts by extrapolation into the forecast year of the trend from the training period?

The ordinary reference forecasts in long range forecasting are trivial, implying extrapolation of climatology from the training period into the forecast year (e.g., Murphy, 1988). For the predictand and predictor(s) series containing essential trend components, it is reasonable to extrapolate the trend from the training period into the forecast year for a trivial reference forecast. Non-climatology based reference forecasts are discussed, e.g., by von Storch and Zwiers (1999).

Results from the analysis shown in Table A2.2 lead to a conclusion that the developed forecast model outperforms the trivial trend forecasts. The PCC and TCC values of the trend based forecasts are essentially less than those of the developed forecasts (Table A2.1). The skill scores of the developed forecasts, with the reference forecasts being those based on the trends, are positive, i.e., the developed forecasts outperform those based on the trends.

Table A2.2 The PCC and TCC of the forecasts based on trends. MSSS and TCCSS of the developed forecasts (dvl.), with the reference forecasts being those based on the trends (tr.).

	PCC	TCC	MSSS dvl./tr.	TCCSS dvl./tr.
March	0.05	0.35	0.25	0.32
April	-0.01	0.14	0.28	0.43
May	0.18	0.47	0.07	0.09

2. What is the contribution of the short term variability component to the performance of the forecasts and what is the contribution of the trend component?

Contribution of the short-term and trend components to co-variability of the observations and developed forecasts may be examined in terms of covariance. These components are not correlated during the training periods, so that each training period predictor value may be reasonably presented as a sum of the trend, and short term variable components and each forecast may be presented as a similar sum corresponding to the forecast year accounting for the corresponding regression coefficient. The series of these forecasts (Fig. A2.1) are not non-correlated. Therefore, the only accurate assessment of their contribution into forecast performance may be formulated in terms of covariance, since the covariance between observations and forecasts may be accurately decomposed into covariance associated with the short term component and trend component. Results (Table A2.3) document that the contribution from the trend exceeds that from the short-term component only for the May forecast verification assessment performed in terms of unadjusted covariance that corresponds to TCC.

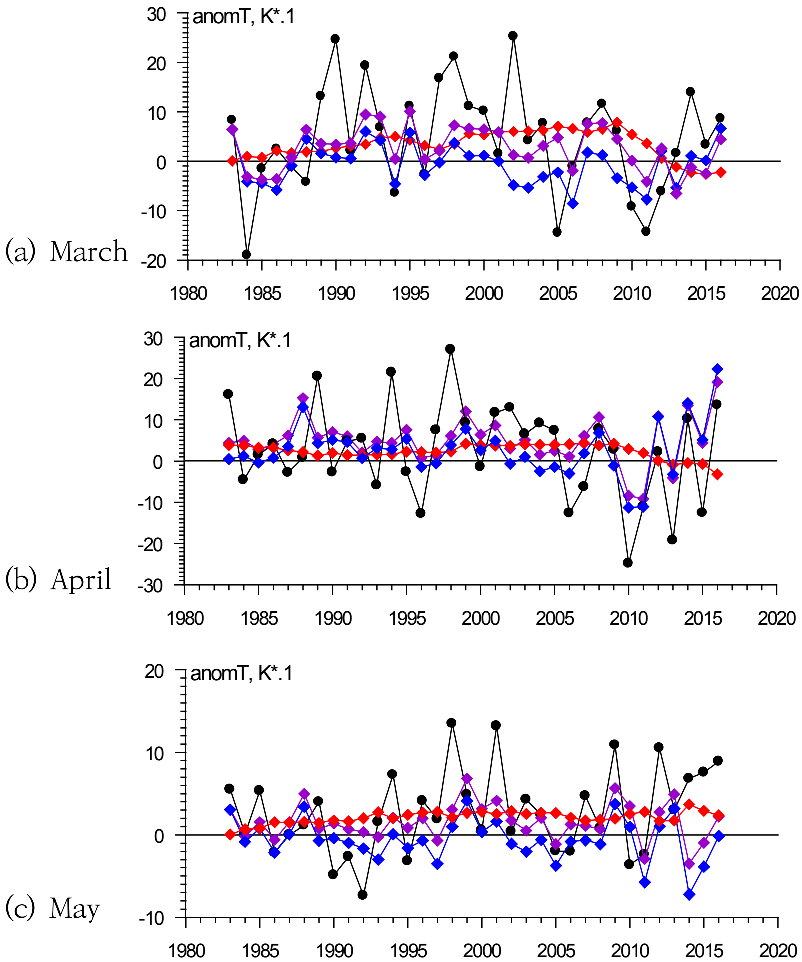


Fig. A2.1 Observed anomalies (black circles), predicted anomalies (purple rhombs), short term (blue rhombs) and trend (red rhombs) components of the predicted anomalies for (a) March, (b) April, and (c) May.

Table A2.3 Covariance between observed and forecast series and its short term variable and trend components. Adjusted covariance corresponds to the PCC, and unadjusted covariance (congruence) corresponds to the TCC, with the PCC and TCC values being shown in the parentheses.

	Total cov (corr)	covariance %	Short cov (corr)	term %	Trend cov (corr)	component %
March						
Adjusted	0.203	100	0.199	98	0.004	2
	(0.45)		(0.46)		(0.01)	
Unadjusted	0.329	100	0.172	53	0.157	47
	(0.56)		(0.36)		(0.31)	
April						
Adjusted	0.325	100	0.316	97	0.009	3
	(0.51)		(0.44)		(0.04)	
Unadjusted	0.447	100	0.380	85	0.067	15
	(0.51)		(0.48)		(0.19)	
May						
Adjusted	0.039	100	0.038	97	0.001	3
	(0.34)		(0.30)		(0.03)	
Unadjusted	0.081	100	0.022	27	0.059	73
	(0.52)		(0.15)		(0.46)	

3. Can the short term component be predictable if the trend does not affect the derivation of the regression equations, i.e., does not exist in all the training periods plus one year?

We show results obtained on the "pure" short term component (in contrast to Point 1 above). For each training period prior to the derivation of the regression equation, we estimated the 25-year trends of both the observations and grid point predictors and removed them from the 26-year series (training period plus forecast year). Verification results of the developed forecasts shown in Table A2.4 document that the developed forecasts are quite skillful, with the deterioration compared to original forecasts being insignificant.

Table A2.4 Verification scores of the forecasts performed on the detrended series, with regression coefficients, not accounting for trends.

	PCC	TCC	MSSS
March	0.43	0.43	0.18
April	0.53	0.51	0.25
May	0.39	0.34	0.10

4. Is there is predictability for the high-pass filtered series?

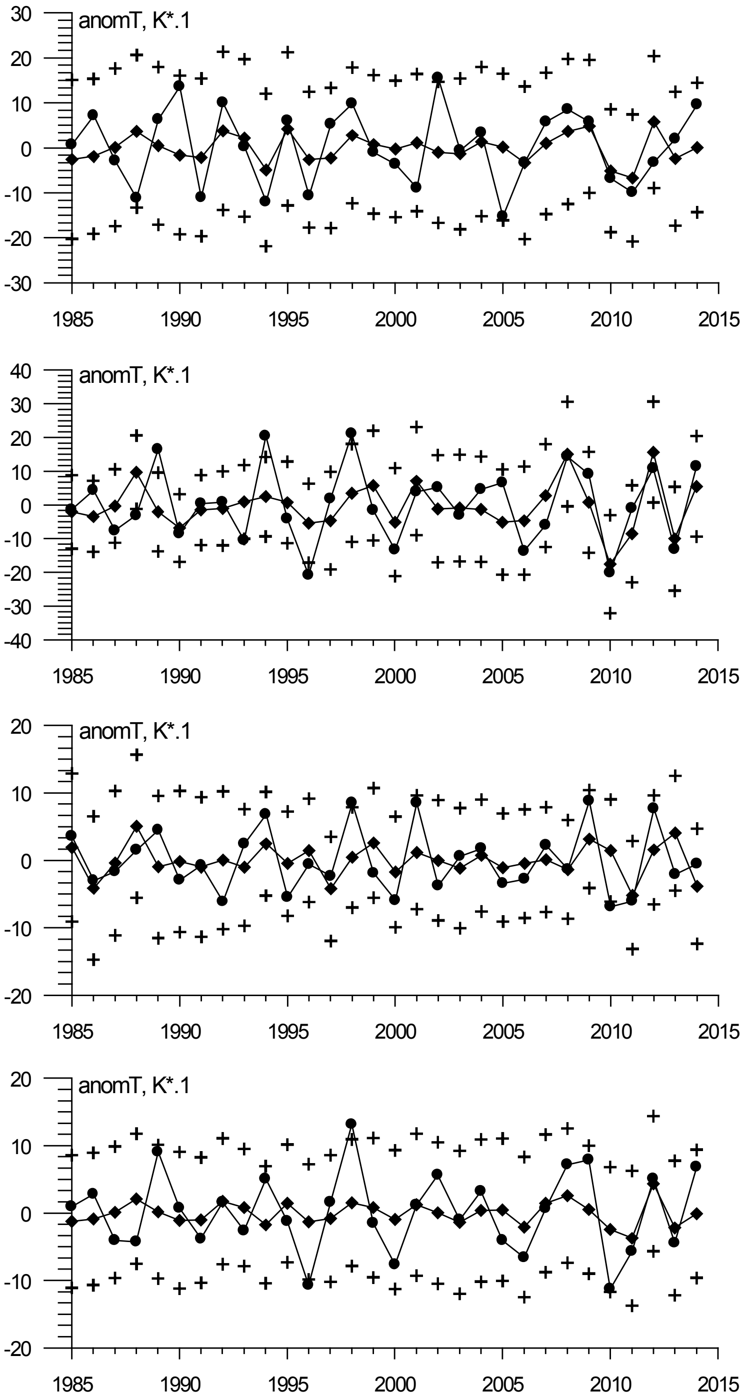
We have retained only the short term components by subtracting running year 3 (5, 7) means from the original series of both observations and grid point predictors. The results are shown in Table A2.5. For all three months and for all three window sizes, the MSSS exceeds 0.10 (significant at the confidence level of 90% and above), and PCC/TCC exceeds 0.33 (significant at the confidence level of 95% and above). It should be noted that the PCC and TCC values are equal with accuracy to the third decimal point, supporting an absence of any trends in the analyzed residual series as well as an absence of any "pillow" in the forecast series.

Table A2.5 Performance of the forecasts of the spring monthly KST anomalies, with the original series being high pass pre-filtered.

Window size: Metric:	3 yrs.		5 yrs.		7 yrs.	
	PCC&TCC	MSSS	PCC&TCC	MSSS	PCC&TCC	MSSS
March	0.39	0.15	0.33	0.10	0.38	0.14
April*	0.39	0.13	0.56	0.31	0.45	0.18
April**	0.55	0.40	0.58	0.33	0.49	0.24
May	0.41	0.16	0.41	0.16	0.48	0.23
MAM	0.53	0.32	0.48	0.32	0.54	0.32

* predictors are Jan. SST and Feb. SST

** predictors are Jan. SST and Feb. SLP



(a) March

(b) April

(c) May

(d) MAM

Fig. A2.2 Series of observed (circles) and forecast (rhombs) anomalies for (a) March, (b) April, (c) May, with the forecasts being performed on the high pass pre-filtered series with the use of the 5-year running mean.

It should be noted that all the verification scores shown in Appendix 2 were obtained on the non-detrended series of the observations and forecasts. Verification scores obtained on the detrended series of the observations and forecasts exceed those shown above.

It may be concluded that the selected predictors for the spring KST anomalies provide reasonable predictability of the spring monthly KST anomalies, particularly of the short-term components of its variations, regardless of the trends in the original series.

APPENDIX 3

Review of the predictors and skill scores of the statistical models developed for prediction of spring Korea temperature.

Prediction from January

predictand		predictor		Skill scores						
Target month	Lead time month	Predictor variable	Area of corr map	TCC	MSSS	corr	RPSS	ROC BN	ROC NN	ROC AN
March	1	SST	45-30N 110-140E	**0.56**	**0.30**	**0.47**	**0.28**	0.69	0.59	**0.84**
April	2	SST	45-20N 110-140E	**0.50**	**0.24**	**0.47**	**0.17**	**0.75**	0.51	0.63
May	*3*	*SST*	*45-0N 110-140E*	**0.42**	*0.14*	*0.17*	*0.07*	*0.45*	*0.52*	*0.57*
MAM	1	SST	45-30N 110-140E	**0.58**	**0.29**	**0.46**	**0.26**	**0.92**	0.42	0.68

Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.

Typed in italic shows that the prediction skill of the model is not appropriate.

Prediction from February

predictand		predictor		Skill scores						
Target month	Lead time month	Predictor variable	Area of corr map	TCC	MSSS	corr	RPSS	ROC BN	ROC NN	ROC AN
April	1	J SST F SST	45-20N 110-140E 40-25N 70-40W	**0.51**	**0.24**	**0.51**	**0.16**	**0.80**	0.36	0.65
April	1	J SST F SLP	45-20N 110-140E 50-40N 80-60W	**0.65**	**0.40**	**0.63**	**0.30**	**0.91**	0.56	**0.75**
May	2	SST	25-30N 110-140E	*0.30*	*0.10*	*0.09*	*0.00*	*0.49*	*0.41*	*0.54*

- Typed in bold is significant at the 5% level (2.5%) in the two (one) tailed test.
- Typed in italic shows that the prediction skill of the model is not appropriate.
- Typed in italic shows that while the prediction skill of the model is high, special caution should be used because of the SLP predictor, which may appear nonstationary.

Prediction from March

predictand		Predictor		Skill scores						
Target month	Lead time month	Predictor variable	Area of corr map	TCC	MSSS	corr	RPSS	ROC BN	ROC NN	ROC AN
May	1	M SST	25-20N 110-140E	**0.52**	**0.27**	**0.34**	0.15	0.62	0.62	0.69

Research Report 2016-13

Development of Physical-Empirical models for seasonal prediction of South Korea spring temperature

Vladimir Kryjov

APEC Climate Center
12 Centum 7-ro, Haeundae-gu, Busan 48058, Republic of Korea
Tel: +82-51-745-3900 Fax: +82-51-745-3949

www.apcc21.org

ISBN 979-11-5698-164-0